



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07K 1/00	A1	(11) International Publication Number: WO 00/14106 (43) International Publication Date: 16 March 2000 (16.03.00)
(21) International Application Number: PCT/US99/21053		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(22) International Filing Date: 9 September 1999 (09.09.99)		
(30) Priority Data: 09/150,489 9 September 1998 (09.09.98) US		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(71) Applicant: THE BURNHAM INSTITUTE [US/US], 10901 N. Torrey Pines Road, La Jolla, CA 92037 (US).		
(72) Inventors: REED, John, C.; 17044 El Camino Real, Rancho Santa Fe, CA 92067 (US). TAKAYAMA, Shinichi, 390 Stratford Court #3, Del Mar, CA 92014 (US).		
(74) Agents: WONG, James, J. et al.; Campbell & Flores LLP, Suite 700, 4370 La Jolla Village Drive, San Diego, CA 92122 (US).		

(54) Title: NOVEL BAG PROTEINS AND NUCLEIC ACID MOLECULES ENCODING THEM

(57) Abstract

The present invention provides a family of BAG-1 related proteins from humans (BAG-1L, BAG-1, BAG-2, BAG-3, BAG-4 and BAG-5), the invertebrate *C. elegans* (BAG-1, BAG-2) and the fission yeast *S. pombe* (BAG-1A, BAG-1B) and the nucleic acid molecules that encode them.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BV	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

NOVEL BAG PROTEINS AND
NUCLEIC ACID MOLECULES ENCODING THEM

STATEMENT AS TO RIGHTS TO INVENTIONS MADE
UNDER FEDERALLY-SPONSORED RESEARCH AND DEVELOPMENT

5 This invention was made with government support under grant number CA-67329 awarded by the National Institutes of Health. The United States Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

10 FIELD OF THE INVENTION

This invention relates generally to the fields of molecular biology and molecular medicine and more specifically to a novel family of proteins that can regulate protein folding. The functions of these proteins 15 are potentially diverse, including promoting tumor cell growth and metastasis.

BACKGROUND INFORMATION

The Hsc70/Hsp70-family of molecular chaperones participate in protein folding reactions, controlling 20 protein bioactivity, degradation, complex assembly/disassembly, and translocation across membranes. These proteins interact with hydrophobic regions within target proteins via a carboxyl (C)-terminal peptide binding domain, with substrate binding and release being controlled 25 by the N-terminal ATP-binding domain of Hsc70/Hsp70. Hsc70/Hsp70-assisted folding reactions are accomplished by repeated cycles of peptide binding, refolding, and release,

which are coupled to ATP hydrolysis by the ATP-binding domain (ATPase) of Hsc70/Hsp70 and by subsequent nucleotide exchange. The chaperone activity of mammalian Hsc70/Hsp70 is regulated by partner proteins that either modulate the 5 peptide binding cycle or that target the actions of these chaperones to specific proteins and subcellular compartments. DnaJ-family proteins (Hdj-1/Hsp40; Hdj-2; Hdj-3) stimulate the ATPase activity of Hsc70/Hsp70, resulting in the ADP-bound state which binds tightly to 10 peptide substrates. The Hip protein collaborates with Hsc70/Hsp70 and DnaJ homologues in stimulating ATP hydrolysis, and thus also stabilize Hsc70/Hsp70 complexes with substrate polypeptides, whereas the Hop protein may provide co-chaperone functions through interactions with 15 the C-terminal peptide binding domain.

The Bcl-2 associated athanogene-1 (bag-1) is named from the Greek word *athanos*, which refers to anti-cell death. BAG-1 was previously referred to as Bcl-2-associated protein-1 (BAP-1) in U.S. Patent No. 20 5,539,094 issued July 23, 1996, which is incorporated herein by reference. In this earlier patent, BAG-1 is described as a portion of the human BAG-1 protein, absent the N-terminal amino acids 1 to 85. In addition, a human protein essentially identical to human BAG-1 was described 25 by Zeiner and Gehring, (*Proc. Natl. Acad. Sci., USA* **92**:11465-11469 (1995)). Subsequent to the issuance of U.S. Patent 5,539,094 the N-terminal amino acid sequence from 1 to 85 of human BAG-1 was reported.

BAG-1 and its longer isoforms BAG-1M (Rap46) and 30 BAG-1L are recently described Hsc70/Hsp70-regulating proteins. BAG-1 competes with Hip for binding to the Hsc70/Hsp70 ATPase domain and promotes substrate release. BAG-1 also reportedly stimulates Hsc70-mediated ATP

hydrolysis by accelerating ADP/ATP exchange, analogous to the prokaryotic GrpE nucleotide exchange protein of the bacterial Hsc70 homologue, DnaK. Gene transfection studies indicate that BAG-1 proteins can influence a wide variety 5 of cellular phenotypes through their interactions with Hsc70/Hsp70, including increasing resistance to apoptosis, promoting cell proliferation, enhancing tumor cell migration and metastasis, and altering transcriptional activity of steroid hormone receptors.

10 Despite the notable progress in the art, there remains an unmet need for the further identification and isolation of additional homologous BAG protein species, and the nucleic acid molecules and/or nucleotide sequences that encode them. Such species would provide additional 15 means by which the identity and composition of the BAG domain, that is, the portion of the protein that is influencing or modulating protein folding, could be identified. In addition, such species would be useful for identifying agents that modulate apoptosis as candidates 20 for therapeutic agents, in particular, anticancer agents. The present invention satisfies these need, as well as providing substantial related advantages.

SUMMARY OF THE INVENTION

The present invention provides a family of BAG-1 25 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO: 4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)] , the invertebrate *C.elegans* [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast *S.pombe* 30 [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)] and the nucleic acid molecules that encode them.

Another aspect of the present invention provides an amino acid sequence present in the family of BAG-1 related proteins, that modulates Hsc70/Hsp70 chaperone activity, that is, the BAG domain.

5 Another aspect of the present invention provides novel polypeptide and nucleic acid compositions and methods useful in modulating Hsc70/Hsp70 chaperone activity.

10 Another aspect of the present invention is directed to methods for detecting agents that modulate the binding of the BAG family of proteins, such as BAG-1 (beginning at residue 116 of SEQ ID NO:2), and related proteins with the Hsc70/Hsp70 Family of proteins or with other proteins that may interact with the BAG-Family proteins.

15 Still another aspect of the present invention is directed to methods for detecting agents that induce the dissociation of a bound complex formed by the association of BAG-Family proteins with Hsc70/Hsp70 Family molecule chaperones or other proteins.

20

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the full length cDNA sequence for human BAG-1 (SEQ ID NO:1) protein with the corresponding amino acid sequence (SEQ ID NO:2). Within the full length sequence are included the overlapping sub-sequences of BAG-1 (beginning at nucleotide 391), BAG-1M [beginning at nucleotide 260 of (SEQ ID NO:2)], and BAG-1L [beginning at nucleotide 46 of (SEQ ID NO:2)].

Figures 2A and 2B combined shows the full length cDNA sequence (SEQ ID NO:3) aligned with the corresponding amino acid residues for human BAG-2 protein (SEQ ID NO:4).

Figure 3 shows a cDNA sequence (SEQ ID NO:5) 5 aligned with the corresponding amino acid residues for human BAG-3 protein (SEQ ID NO:6).

Figure 4 shows the a cDNA sequence (SEQ ID NO:7) aligned with the corresponding amino acid residues for human BAG-4 protein (SEQ ID NO:8).

10 Figure 5 shows a cDNA sequence (SEQ ID NO:9) aligned with the corresponding amino acid residues for human BAG-5 protein (SEQ ID NO:10).

Figure 6A shows the full length cDNA sequence for *C. elegans* BAG-1 protein (SEQ ID NO:11).

15 Figure 6B shows the 210 amino acid sequence for *C. elegans* BAG-1 protein (SEQ ID NO:12).

Figure 7A shows the full length cDNA sequence for *C. elegans* BAG-2 protein (SEQ ID NO:13).

20 Figure 7B shows the 458 amino acid sequence for *C. elegans* BAG-2 protein (SEQ ID NO:14).

Figure 8A shows the full length cDNA sequence for *S. pombe* BAG-1A protein (SEQ ID NO:15).

Figure 8B shows the 195 amino acid sequence for *S. pombe* BAG-1A protein (SEQ ID NO:16).

Figure 9A shows the full length cDNA sequence for *S. pombe* BAG-1B protein (SEQ ID NO:17).

Figure 9B shows the 206 amino acid sequence for *S. pombe* BAG-1B protein (SEQ ID NO:18).

5 Figure 10 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10); *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). (A) The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like nuclear localization sequence are also shown. (B) The amino acid 10 sequences of the BAG domain for human BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10), *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18), and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) are aligned demonstrating 15 their homology. Black and gray shading represent identical 20 and similar amino acids, respectively.

Figure 11 shows assays demonstrating the interaction of BAG-family proteins with Hsc70/ATPase. (A) Two-hybrid assays using yeast expressing the indicated 25 fusion proteins. Blue color indicates a positive interaction, resulting in activation of the *lacZ* reporter gene. (B) *In vitro* protein assays using GST-fusion proteins and 35 S-labeled *in vitro* translated proteins. (C) Co-immunoprecipitation assays using anti-Flag or IgG1 30 control antibodies and lysates from 293T cells expressing Flag-tagged BAG-1 (beginning at residue 116 of SEQ ID

NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), Daxx, or Apaf-1.

Figure 12 shows surface plasmon resonance analysis of BAG-family protein interactions with 5 Hsc70/ATPase. (A) SDS-PAGE analysis of purified recombinant proteins. (B) Representative SPR results of biosensor chips containing immobilized BAG proteins with and without maximally bound Hsc70/ATPase.

Figure 13 shows representative SPR results for 10 biosensor chips containing immobilized BAG-1 (beginning at residue 116 at SEQ ID NO:2), BAG-1(ΔC), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) proteins. Hsc70/ATPase was flowed over the chips (arrow/left) until maximal binding was reached (response units), then flow was continued without 15 Hsc70/ATPase (arrow/right). For BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6), Hsc70 was injected at 0.0175, 0.035, 0.07, 0.14, and 0.28 μ M.

Figure 14 shows BAG-family protein modulation of Hsc70 chaperone activity. (A) Protein refolding assay of 20 chemically-denatured luciferase by Hsc70 plus DnaJ in the absence or presence of BAG and BAG-mutant proteins. (B) Concentration-dependent inhibition of Hsc70-mediated protein refolding by BAG-family proteins [BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 25 (SEQ ID NO:6)] but not by BAG-mutant (BAG-1 (ΔC)). (C) Hsc70/Hsp40-mediated refolding of heat-denatured luciferase was assayed in the presence of (black bars) or absence of (striped bars) of 1.8 μ M Hip, with (lanes 3-10) or without (lanes 1,2) various BAG-family proteins (1.8 μ M) as 30 indicated (mean \pm SE; n=3). A control (CNTL) is shown (lane 1) in which Hsc70 was replaced with an equivalent amount of BSA.

Figure 15A shows an expanded cDNA sequence for human BAG-3 protein (SEQ ID NO:19).

Figure 15B shows the corresponding amino acid residues for the human BAG-3 protein (SEQ ID NO:20) of 5 Figure 15A.

Figure 15C shows the expanded cDNA sequence (SEQ ID NO:19) aligned with the corresponding amino acid residues for human BAG-3 protein of Figure 15A (SEQ ID NO:20).

10 Figure 16A shows an expanded cDNA sequence for human BAG-4 protein (SEQ ID NO:21).

Figure 16B shows the corresponding amino acid residues for the human BAG-4 protein of Figure 16A (SEQ ID NO:22).

15 Figure 16C shows the expanded cDNA sequence (SEQ ID NO:21) aligned with the corresponding amino acid residues for human BAG-4 protein of Figure 16A (SEQ ID NO:22).

Figure 17A shows an expanded cDNA sequence for 20 human BAG-5 protein (SEQ ID NO:23).

Figure 17B shows the corresponding amino acid residues for the human BAG-5 protein of Figure 17A (SEQ ID NO:24).

25 Figure 17C shows the expanded cDNA sequence (SEQ ID NO:23) aligned with the corresponding amino acid residues for human BAG-5 protein of Figure 17A (SEQ ID NO:24).

Figure 18 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), expanded BAG-3 (SEQ ID NO:20), expanded BAG-4 (SEQ ID NO:22), expanded BAG-5 (SEQ ID NO:24); 5 *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like 10 nuclear localization sequence are also shown.

Definitions

The term "apoptosis", as used herein, refers to the process of programmed cell death, although not all programmed cell deaths occur through apoptosis, as used 15 herein, "apoptosis" and "programmed cell death" are used interchangeably.

The term "tumor cell proliferation", as used herein refers to the ability of tumor cells to grow and thus expand a tumor mass.

20 The term "cell migration", as used herein refers to the role cell motility plays in the invasion and potentially metastasis by tumor cells.

The term "metastasis", as used herein refers to the spread of a disease process from one part of the body 25 to another, as in the appearance of neoplasms in parts of the body remote from the site of the primary tumor; results in dissemination of tumor cells by the lymphatics or blood vessels or by direct extension through serous cavities or subarachnoid or other spaces.

The term "steroid hormone receptor function", as used herein refers to physiological, cellular and molecular functioning of receptors sites that bind with steroid hormones.

5 The term "substantially purified", as used herein, refers to nucleic acid or amino acid sequence that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with
10 which they are naturally associated.

"Nucleic acid molecule" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single or double stranded,
15 and represent the sense or antisense strand.

"Hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

20 The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T binds to the complementary sequence "T-C-A".

25 The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid and is referred to using the functional term
30 "substantially homologous." The inhibition of

hybridization of the completely complementary sequence to the target sequence may be examined using a hybridzation assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A 5 substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency.

The term "antisense", as used herein, refers to 10 nucleotide sequences which are commplementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of 15 interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further 20 transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense, and "positive" is sometimes used in reference to the sense strand.

25 "Amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited herein this term excludes an amino 30 acid sequence of a naturally occurring protein. "Amino acid sequence", "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

The term "functional fragments" or "fragments", as used herein, with regard to a protein refers to portions of that protein that are capable of exhibiting or carrying out the activity exhibited by the protein as a whole. The 5 portions may range in size from three amino acid residues to the entire amino acid sequence minus one amino acid. For example, a protein "comprising at least a functional fragment of the amino acid sequence of SEQ ID NO:1", encompasses the full-length of the protein of SEQ ID NO:1 10 and portions thereof.

A "derivative" of a BAG protein, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The derivative may have "conservative" changes, wherein a substituted amino acid has similar 15 structural or chemical properties, e.g., substitution of an apolar amino acid with another apolar amino acid (such as replacement of leucine with isoleucine). The derivative may also have "nonconservative" changes, wherein a substituted amino acid has different but sufficiently 20 similar structural or chemical properties that permits such a substitution without adversely effecting the desired biological activity, e.g., replacement of an amino acid with an uncharged polar R group with an amino acid with an apolar R group (such as replacement of glycine with 25 tryptophan), or alternatively replacement of an amino acid with a charged R group with an amino acid with an uncharged Polar R group (such as replacement of lysine with asparagine).

Amino Acids - Apolar R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
alanine	methyl	ala	A
valine	2-propyl	aal	V
leucine	2-methylpropyl	leu	L
isoleucine	2-butyl	ile	I
proline	propyl* - cyclized	pro	P
phenylalanine	benzyl	phe	F
tryptophan	3-indolylmethyl	tyr	W
methionine	methylthioethyl	met	M

Amino Acids - Uncharged Polar R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
glycine	H	gly	G
serine	hydroxymethyl	ser	S
threonine	1-hydroxyethyl	thr	T
cysteine	thiolmethyl	cys	C
tyrosine	4-hydroxyphenylmethyl	tyr	Y
asparagine	aminocarbonylmethyl	asn	N
glutamine	aminocarbonylethyl	gln	Q

20 Amino Acids - Charged R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
aspartic acid	carboxymethyl	asp	D
glutamic acid	carboxyethyl	glu	E
lysine	4-aminobutyl	lys	K
arginine	3-guanylpropyl	arg	R
histidine	4-imidazoylmethyl	his	H

Similar minor modifications may also include amino acids deletions or insertions or both. Guidance in determining which amino acid residues may be modified as indicated above without abolishing the desired biological 5 functionality may be determined using computer programs well known in the art, for example, DNASTAR software. In addition, the derivative may also result from chemical modifications to the encoded polypeptide, including but not limited to the following, replacement of hydrogen by an 10 alkyl, acyl, or amino group; esterification of a carboxyl group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative. Further a derivative may also result from the substitution of a L- configuration amino acid with its corresponding D- 15 configuration counterpart.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of a protein/polypeptide or portions thereof (such as BAG-1) and, as such, is able to 20 effect some or all of the actions of BAG-1 protein.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene 25 agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al., *Anticancer Drug Des.* 8:53-63 (1993)).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a family of BAG-1 30 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1S beginning at residue 116 of SEQ ID NO:2, BAG-2 (SEQ ID

NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO: 8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate *C.elegans* [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast *S.pombe* 5 [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)], specifically the full length amino acid sequences comprising human BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), and BAG-2 (SEQ ID NO:4) *C. elegans* BAG-1 (SEQ ID NO:12), and BAG-2 (SEQ ID NO:14), and 10 *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and partial sequences comprising human BAG-3 (SEQ ID NO: 6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) and functional fragments thereof. In particular, the invention provides 15 the amino acid sequences comprising human BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) proteins.

Another aspect of the present invention provides 20 the nucleic molecule and nucleotide sequences that encode the family of BAG-1 related proteins from humans [BAG-1 (SEQ ID NO:1), BAG-2 (SEQ ID NO:3), BAG-3 (SEQ ID NO:5) and (SEQ ID NO:19), BAG-4 (SEQ ID NO:7) and (SEQ ID NO:21) and BAG-5 (SEQ ID NO:9) and (SEQ ID NO:23)], the invertebrate 25 *C.elegans* [BAG-1 (SEQ ID NO:11), BAG-2 (SEQ ID NO:13)] and the fission yeast *S.pombe* [BAG-1A (SEQ ID NO:15), BAG-1B (SEQ ID NO:17)].

BAG-1L (SEQ ID NO:2) is a multifunctional protein that blocks apoptosis, promotes tumor cell metastasis, and 30 contributes to factor-independent and p53-resistant cell growth. BAG-1L (SEQ ID NO:2) interacts with several types of proteins, including Bcl-2, some tyrosine kinase growth

factor receptors, steroid hormone receptors, and the p53-induced cell cycle regulator Siah-1A.

BAG-1 is a regulator of Hsc70/Hsp70 family molecular chaperones. A carboxyl-terminal domain in this 5 protein binds tightly to the ATPase domains of Hsc70 and Hsp70 ($K_D = 1$ nM) (Zeiner, M., Gebauer, M., and Gehring, U., *EMBO J.* **16**: 5483-5490, (1997)). BAG-1 modulates the activity of these molecular chaperones, acting as an apparent functional antagonist of the Hsp70/Hsc70-10 associated protein Hip (3-5) (Höhfeld, J. and Jentsch, S., *EMBO J.* **16**: 6209-6216, (1997); Takayama, S., Bimston, D. N., Matsuzawa, S., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R. J., and Reed, J. C., *EMBO J.* **16**: 4887-96, (1997); Zeiner, M., Gebauer, M., and Gehring, U., *EMBO J.* 15 **16**: 5483-5490, (1997)). In general, protein refolding is accomplished by Hsp70/Hsc70 through repeated cycles of target peptide binding and release, coupled to ATP hydrolysis (Ellis, R., *Curr Biol.* **7**: R531-R533, (1997)). BAG-1 appears to promote substrate release, whereas Hip 20 stabilizes Hsp70/Hsc70 complex formation with target peptides (Höhfeld, J., Minami, Y., and Hartl, F.-U., *Cell.* **83**: 589-598, (1995)). Since each substrate interaction with Hsc70/Hsp70 is unique in terms of the optimal length of time the protein target should remain complexed with 25 Hsc70/Hsp70 for achieving new conformations, the net effect of BAG-1 can be either enhancement or inhibition of the refolding reaction.

The 70kd heat shock family proteins (Hsp70/Hsc70) are essential to a variety of cellular processes and have 30 been implicated in cancer, yet it is unclear how these proteins are regulated *in vivo*. A variety of co-chaperones have been identified which may target Hsp70/Hsc70 to different subcellular compartments or promote their

interactions with specific protein or protein complexes. BAG-1 appears to represent a novel Hsp70/Hsc70 regulator which differs functionally from all other mammalian co-chaperones identified to date, such as members of the 5 DnaJ-, Hip-, Hop-, and cyclophilin-families of proteins.

Another aspect of the present invention provides the amino acid sequence of a binding domain of about 40 to 55 amino acids that bind the a Hsc70/Hsp70 ATPase domain. The BAG domain is situated near the C-terminus, and the 10 ubiquitin-like domains are situated near the N-terminus.

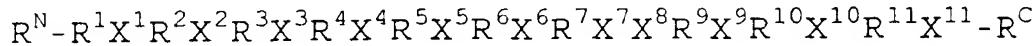
The BAG family of proteins of the present invention contain a common conserved C-terminal domain (the "BAG" domain) that facilitates binding to the ATPase domain of Hsp70/Hsc70. The carboxyl-terminal domain of BAG-1 15 binds to the ATPase domain of Hsc70/Hsp70 and regulates its chaperone function by acting as a ADP-ATP exchange factor. Other domains of BAG-1 mediate interactions with proteins such as Bcl-2 and retinoic acid receptors (RARs), allowing BAG-1 to target Hsc70/Hsp70 to other proteins, presumably 20 modulating their function by changing their conformations.

Human BAG-1 was previously shown to inhibit Hsc70/Hsp70 dependent refolding of denatured protein substrates *in vitro* (S. Takayama, et al., *EMBO J* 16, 4887-96 (1997); M. Zeiner, M. Gebauer, U. Gehring, *EMBO J.* 16, 25 5483-5490 (1997); and J. Höhfeld, S. Jentsch, *EMBO J.* 16, 6209-6216 (1997)). In Example III, Part A the effects of recombinant human BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were compared using *in vitro* protein refolding assays similar to those employed previously for assessing 30 BAG-1. The study showed that addition of equimolar amounts of each of the recombinant proteins to Hsc70 resulted in significant inhibition of luciferase refolding, with BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) showing somewhat

greater inhibitor activity than BAG-1 (Figure 4A). In a separate luciferase folding study BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) once again displayed inhibition of luciferase refolding, however in this study 5 varying amounts of BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were added relative to Hsc70 which resulting in concentration-dependent inhibition of Hsc70 chaperone activity, i.e., luciferase folding (Example III Part A). Additional follow on studies using the same experimental 10 protocols as the previous studies, as taught in Example IIA, have shown that BAG-4 (SEQ ID NO:22) also undergoes association with Hsc70/ATPase.

Yet another aspect of the present invention provides a nucleotide sequence having at least about 15 15 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides that can hybridize or is complementary under relatively stringent conditions to a portion of the nucleic acid 20 sequences shown in Figures 1-9 and Figures 15-17, in particular the BAG domain as shown in Figure 1B, e.g., nucleotides 552-593 of human BAG-3, or nucleotides 167-221 of human BAG-4.

Yet another aspect of the present invention 25 provides a compound of the formula,



wherein,

R^N is a group of 1 to 552 independently selected amino acids;

30 R^1 is a group of 3 independently selected amino acids;

X^1 is an amino acid with a charged or uncharged R group, such as aspartic acid, glutamic acid, asparagine, or glutamine;

5 R^2 is a group of 7 independently selected amino acids;

X^2 is an amino acid with a charged R group, such as glutamic acid;

R^3 is a group of 5 independently selected amino acids;

10 X^3 is an amino acid with an apolar R group, such as leucine, methionine, or isoleucine;

R^4 is a group of 3 independently selected amino acids;

15 X^4 is an amino acid with charged R group, such as aspartic acid or glutamine acid;

R^5 is a single independently selected amino acid;

X^5 is an amino acid with apolar or uncharged R group, such as leucine, valine, methionine, alanine or threonine;

20 R^6 is a group of 15 independently selected amino acids;

X^6 is an amino acid with a charged or uncharged R group, such as arginine, lysine, glutamine or aspartic acid;

25 R^7 is a group of 2 independently selected amino acids;

X^7 is an amino acid with a charged R group, such as arginine;

30 X^8 is an amino acid with a charged R group, such as arginine or lysine;

R^9 is a group of 2 independently selected amino acids;

X^9 is an amino acid with an apolar R group, such as valine;

35 R^{10} is a group of 3 independently selected amino acids;

X^{10} is an amino acid with an uncharged R group, such as glutamine;

R^{11} is a group of 2 independently selected amino acids;

5 X^{11} is an amino acid with an apolar R group, such as leucine; and

R^c is a group of 1 to 100 independently selected amino acids.

A nucleotide sequence of at least about 15
10 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides can be useful, for example, as a primer for the polymerase chain reaction (PCR) or other similar reaction mediated by
15 a polymerase such as a DNA or RNA polymerase (see PCR Protocols: A guide to methods and applications, ed. Innis et al. (Academic Press, Inc., 1990), which is incorporated herein by reference; see, for example, pages 40-41). In addition, such a nucleotide sequence of the invention can
20 be useful as a probe in a hybridization reaction such as Southern or northern blot analysis or in a binding assay such as a gel shift assay.

A nucleotide sequence of the invention can be particularly useful as an antisense molecule, which can be
25 DNA or RNA and can be targeted to all or a portion of the 5'-untranslated region or of the 5'-translated region of a bag-1 nucleic acid sequence in a cell. For example, an antisense molecule can be directed to at least a portion of the sequence shown as the BAG domain in Figure 1A, e.g.,
30 nucleotides 272-319 of human BAG-1L (SEQ ID NO:1), or nucleotides 79-147 of human BAG-5 (SEQ ID NO:9). Since the 5'-region of a nucleic acid contains elements involved in the control of expression of an encoded protein, an antisense molecule directed to the 5'-region of a nucleic

acid molecule can affect the levels of protein expressed in a cell.

A nucleotide sequence of the invention also can be useful as a probe to identify a genetic defect due a 5 mutation of a gene encoding a BAG protein in a cell. Such a genetic defect can lead to aberrant expression of a BAG protein in the cell or to expression of an aberrant BAG protein, which does not properly associate with a Bcl-2-related protein or Hsc70/Hsp70 protein in the cell. As a 10 result, a genetic defect in a gene encoding, for example, human BAG-1 can result in a pathology characterized by increased or decreased levels in protein folding.

Further a nucleotide compound or composition as taught in the present invention can be synthesized using 15 routine methods or can be purchased from a commercial source. In addition, a population of such nucleotide sequences can be prepared by restriction endonuclease or mild DNase digestion of a nucleic acid molecule that contains nucleotides as shown in the nucleotide sequences 20 shown in Figures 1-9 and Figures 15-17 that encodes the amino acids sequences also shown in Figures 1-9 and Figures 15-17. Methods for preparing and using such nucleotide sequences, for example, as hybridization probes to screen a library for homologous nucleic acid molecules 25 are well known in the art (see, for example, Sambrook et al., *Molecular Cloning: A laboratory manual* (Cold Spring Harbor Laboratory Press 1989); Ausubel et al., *Current Protocols in Molecular Biology* (Green Publ., NY 1989), each of which is incorporated herein by reference).

30 A particular nucleotide sequence can be designed based, for example, on a comparison of the nucleic acid molecules encoding any one of the BAG family proteins, as shown in Figures 1-9 and Figures 15-17, with another in the family. Such a comparison allows, for example, the

preparation of a nucleotide sequence that will hybridize to a conserved region present in both nucleic acid molecules, thus providing a means to identify homologous nucleic acid molecules present in other cell types or other organisms.

5 In addition, such a comparison allows the preparation of a nucleotide sequence that will hybridize to a unique region of any of the BAG family nucleotide sequences, such as those corresponding to the BAG domain, thus allowing identification of other proteins sharing this motif. In

10 this regard, it is recognized that, while the human BAG-3 proteins shown as Figures 3 and 20, and human BAG-5 proteins shown as Figures 5 and 24, are only partial sequences, a variant human BAG-3 or BAG-5 produced, for example, by alternative splicing can exist and can be

15 identified using an appropriately designed nucleotide sequence of the invention as a probe. Such useful probes readily can be identified by inspection of the sequences shown in the disclosed Figures by a comparison of the encoding nucleotide sequences.

20 If desired, a nucleotide sequence of the invention can incorporate a detectable moiety such as a radiolabel, a fluorochrome, a ferromagnetic substance, a luminescent tag or a detectable binding agent such as biotin. These and other detectable moieties and methods of

25 incorporating such moieties into a nucleotide sequence are well known in the art and are commercially available. A population of labelled nucleotide sequences can be prepared, for example, by nick translation of a nucleic acid molecule of the invention (Sambrook et al., *supra*,

30 1989; Ausubel et al., *supra*, 1989).

One skilled in the art would know that a method involving hybridization of a nucleotide sequence of the invention can require that hybridization be performed under relatively stringent conditions such that nonspecific

35 background hybridization is minimized. Such hybridization

conditions can be determined empirically or can be estimated based, for example, on the relative GC content of a sequence and the number of mismatches, if known, between the probe and the target sequence (see, for example, 5 Sambrook et al., *supra*, 1989).

The invention further provides antibodies specific for human BAG family protein. As used herein, the term "antibody" includes polyclonal and monoclonal antibodies, as well as polypeptide fragments of antibodies 10 that retain a specific binding activity for human BAG-1 of at least about 1×10^5 M⁻¹. One skilled in the art would know that anti-BAG-1 antibody fragments such as Fab, F(ab'), and Fv fragments can retain specific binding activity for human BAG-1 (beginning at residue 116 of SEQ ID NO:2) and, 15 thus, are included within the definition of an antibody. In addition, the term "antibody" as used herein includes naturally occurring antibodies as well as non-naturally occurring antibodies and fragments that retain binding activity such as chimeric antibodies or humanized 20 antibodies. Such non-naturally occurring antibodies can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains as described by Huse 25 et al., *Science* 246:1275-1281 (1989), which is incorporated herein by reference.

One skilled in the art would know that purified BAG family protein, which can be prepared from natural sources or synthesized chemically or produced 30 recombinantly, or portions of a BAG family protein, including a portion of human BAG family protein such as a synthetic peptide as described above, can be used as an immunogen. Such peptides useful for raising an antibody include, for example, peptide portions of the N-terminal 85 35 amino acids or the BAG domain of any of the human BAG

proteins (see Figure 1B). A particularly advantageous use of such a protein is for the immunostaining, wherein the methods provides a process to contrast the immunostaining of BAG-family proteins in carcinoma cells with adjacent 5 non-neoplastic prostatic epithelial and basal cells which are generally present in the same tissue sections. These results would be correlated with a Gleason grade to determine whether any of the BAG-family proteins tend to be expressed at higher or lower levels in histologically 10 advanced tumors. From this process a determination can be made as to degree at which the disease is progressing in a given patient, i.e., a prognosis can be made.

Non-immunogenic fragments or synthetic peptides of BAG proteins can be made immunogenic by coupling the 15 hapten to a carrier molecule such bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH), as described in Example IV, below. In addition, various other carrier molecules and methods for coupling a hapten to a carrier molecule are well known in the art and described, for 20 example, by Harlow and Lane, *Antibodies: A laboratory manual* (Cold Spring Harbor Laboratory Press, 1988), which is incorporated herein by reference.

EXAMPLES

The following examples are given to enable those 25 skilled in the art to more clearly understand and to practice the present invention. They should not be considered as limiting the scope of the invention, but merely as being illustrative and representative thereof.

EXAMPLE I

Isolation and Characterization
of BAG-family cDNA Sequences

This example describes methods for isolating and
5 characterizing of BAG-family cDNA sequences from human,
nematode and yeast.

A. Cloning of human BAG cDNA sequences

Yeast two-hybrid library screening of a human
Jurkat cell cDNA library was performed as described by
10 Takayama et al., EMBO J. 16:4887-96 (1997); Matsuzawa et
al., EMBO J. 17:2736-2747 (1998), which are incorporated
herein by reference) using EGY48 strain yeast transformed
with pGilda-Hsc70/ATPase (67-377 amino acids) and the lacZ
reporter plasmid pSH18-34. Of the resulting $\sim 5 \times 10^6$
15 transformants, 112 Leu^r colonies were obtained after
1 week incubation at 30°C. Assay of β -galactosidase (β -gal)
activity of these colonies resulted in 96 clones. Mating
tests were then performed using RFY206 yeast strain
transformed with pGilda, pGilda mBAG-1 (1-219), or pGilda
20 Hsc70/ATPase. Of these, 66 displayed specific interactions
with Hsc70/ATPase. The pJG4-5 cDNAs were recovered using
KC8 *E. coli* strain which is auxotrophic for tryptophan
(Trp). DNA sequencing revealed 3 partially overlapping
25 human BAG-1, 4 identical and one overlapping cDNAs encoding
BAG-2, and 2 partially overlapping BAG-3 clones.

Using the above described yeast two-hybrid screen
with the ATPase domain of Hsc70 as "bait", several human
cDNAs were cloned which encode portions of BAG-1 or of two
other BAG-1-like proteins which are termed BAG-2 (SEQ ID
30 NO:4) and BAG-3 (SEQ ID NO:6). The longest of the cDNAs
for BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) contained
open reading frames (ORFs) of 207 and 162 amino acids,
respectively, followed by stop codons. All BAG-1 (SEQ ID

NO:1), BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) cDNAs obtained by two-hybrid library screening with Hsc70/ATPase contained a conserved domain of about 40-50 amino acids which are termed the "BAG" domain and are shown in Figure 5 10. These results demonstrate that a family of BAG-1-related proteins all contain a conserved ~45 amino acid region near their C-terminus that binds Hsc70/Hsp70.

B. Identification of additional BAG-family proteins

A search of the translated Genbank database using 10 the bBLAST and FASTA search programs also identified human ESTs that provided sequences for further investigation of BAG-family proteins. The putative BAG-4 (SEQ ID NO:8) and BAG-5 (SEQ ID NO:10) proteins contain BAG-domains that share the greatest sequence similarity with the BAG-domain 15 of BAG-3 (SEQ ID NO:6). These were designated BAG-4 (Accession number AA693697, N74588) and BAG-5 (Accession number AA456862, N34101). BAG-4 has 62% identity and ~81% similarity to BAG-3, and BAG-5 has 51% identity and ~75% similarity to BAG-3.

20 Additional BAG-family orthologues or homologues were also identified using computer-based searches and resulted in BAG-family homologue in the nematode *C. elegans* and the fission yeast *S. pombe*. The *C. elegans* genome encodes two apparent BAG-family proteins, which are most 25 similar in their overall sequences to the human BAG-1 (Afo39713, gi:2773211) (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) (Afo68719, gi:3168927). The *S. pombe* contains two BAG-family proteins that share the greatest overall sequence similarity with human BAG-1 (Alo23S54, gi/3133105 30 and Alo23634, gi/3150250). The human and *C. elegans* BAG-1 proteins as well as *S. pombe* BAG-1A all have ubiquitin-like domains near their N-termini (see Figure 10A) of unknown function.

The overall predicted amino acid sequences of the *C. elegans* BAG-1 (SEQ ID NO:12) and *S. pombe* BAG-1A (SEQ ID NO:16) proteins are ~18% identical (~61% similar) and ~17% identical (~64% similar), respectively, to human BAG-1, 5 implying origin from a common ancestral gene. The *C. elegans* BAG-1 protein (SEQ ID NO:12), however, contains a 5 to 7 amino acid insert in its BAG-domain as compared to the human, murine, and yeast BAG-1 homologues (see Figure 10B), and is more similar to BAG-2 (SEQ ID NO:4) in regard 10 to its BAG-domain. *C. elegans* and human BAG-2 also may be derived from a common ancestor as the C-terminal 225 amino acid region which encompasses both the BAG domain and upstream region of both *C. elegans* and human BAG-2 share ~34% amino acid sequence identity and ~70% similarity. The 15 human BAG-2 protein (SEQ ID NO:4), however, contains a 9 amino acid insert in its BAG-domain compared to its *C. elegans* counterpart (see Figure 10B). Evolutionary-tree prediction algorithms suggest that human and *C. elegans* BAG-2 represent a distinct branch of the BAG-family that is 20 more evolutionarily distant from the other BAG-family proteins. None of the predicted BAG-family proteins contain recognizable regions analogous to those found in other Hsc70 regulatory proteins, such as the J-domains and G/F-domains of DnaJ family proteins and the 25 Tetratricopeptide Repeat (TR) domains of Hip/Hop family proteins.

C. Yeast two-hybrid assay of BAG binding to Hsc70/ATPase

The longest of the cDNAs obtained for the BAG-2 and BAG-3 proteins were expressed with N-terminal 30 transactivation (TA) domains in yeast and tested by yeast two-hybrid assay for interactions with fusion proteins consisting of Hsp70/ATPase or a variety of unrelated proteins (Fas, Siah, Fadd) containing N-terminal LexA DNA-binding domains. TA-BAG-2 and TA-BAG-3 demonstrated

positive interactions with LexA-Hsc70/ATPase, resulting in transactivation of a lacZ reporter gene that was under the control of LexA operators (Figure 11A). No interactions with LexA-Fas (cytosolic domain), LexA-Siah, LexA-Fadd, or 5 LexA were detected (see Figure 11A) demonstrating that the BAG-2 and BAG-3 proteins interact specifically with Hsc70/ATPase. Specific two-hybrid interactions between Hsc70/ATPase and either BAG-2 or BAG-3 were also observed when BAG-2 and BAG-3 were expressed as LexA DNA-binding 10 domain fusion proteins and Hsc70/ATPase was fused with a TA domain (see Figure 11A; right panel). These results demonstrate that similarly to BAG-1, BAG-2 and BAG-3 specifically interact with Hsc70/ATPase.

In order to determine whether the BAG proteins 15 are capable of forming heterodimers, coexpression of BAG-2 and BAG-3 in the yeast two-hybrid assay was also performed. Coexpression of BAG-2 and BAG-3 failed to show interaction with BAG-1 or a deletion mutant of BAG-1 (Δ C) which is missing part of its C-terminal domain required for 20 Hsp70/Hsc70 binding suggest that these proteins do not form heterodimers.

D. Isolation and characterization of the complete open reading frame sequences of BAG-2 and BAG-3

In order to deduce the complete ORFs of BAG-2 and 25 BAG-3, a λ -phage cDNA library was screened as follows, using hybridization probes derived from the two-hybrid screening. A human jurkat T-cell λ -ZapII library cDNA library (Stratagene) was screened by hybridization using 32 P-labeled purified insert DNA from the longest of the 30 human BAG-2 (clone #11) and human BAG-3 (clone #28) cDNA clones. From about one million clones screened, 38 BAG-2 and 23 BAG-3 clones were identified, cloned, and their cDNA inserts recovered as pSKII plasmids using a helper phage method (Stratagene). DNA sequencing of λ -phage derived

human BAG-2 cDNA clones revealed an ORF encoding a predicted 211 amino acid protein, preceded by an in-frame stop codon. The longest human BAG-3 λ -phage cDNA clone contains a continuous ORF of 682 amino acids followed by a 5 stop codon, but without an identifiable start codon (see Figure 10A).

Although BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) all contain a homologous BAG domain 10 near their C-terminus, the N-terminal regions of these proteins are dissimilar. Using a combination of search tools (Prosite Search: PP search, using the Prosite pattern database, BCM Search Launcher, Baylor College of Medicine, and Blocks Search), it was determined that the BAG-2 N- 15 terminal region contains potential kinase phosphorylation sites but otherwise shares no apparent similarity with other proteins or known functional domains.

In contrast, the predicted N-terminal region BAG-3 contains a WW domain as shown in Figure 10A. WW 20 domains have been identified in a wide variety of signaling proteins, including a Yes kinase adaptor protein (YAP), the Na⁺-channel regulator Nedd4, formin-binding proteins, dystrophin, and the peptidyl prolyl cis-trans-isomerase Pin-1. These roughly 40 amino acid domains mediate protein 25 interactions and bind the preferred peptide ligand sequence xPPxY (Sudol., TIBS, 21: 161-163, 1996, which is incorporated herein by reference).

EXAMPLE IIIn vitro Association of
BAG proteins and Hsc70/ATPase

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) bind Hsc70/ATPase in various *in vitro* assays.

A. Solution binding assay of BAG-2 and BAG-3 to
Hsc70/ATPase

Association of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) with Hsc70/ATPase was determined by an *in vitro* protein binding assay where Hsc70/ATPase or BAG-family proteins were expressed in bacteria as Glutathione S-Transferase (GST) fusion proteins. Purified cDNA sequences encoding residues 5 to 211 of human BAG-2 (clone #11) and the C-terminal 135 amino acids of human BAG-3 (clone #28) (see Figure 10A) were subcloned into the EcoRI/Xho I sites of pGEX4T-1 prokaryotic expression plasmid (Pharmacia; Piscataway, NJ). These plasmids as well as pGEX4T-1-BAG-1, pGEX-4T-1-BAG-1 (Δ C), and pGEX-4T-1-XL which have been described previously (Takayama et al., *supra* (1997); Xie et al., *Biochemistry*, 37:6410-6418, (1998), which are incorporated herein by reference), were expressed in XL-1 blue strain *E. Coli* (Stratagene, Inc., La Jolla, CA). Briefly, a single colony was inoculated into 1L of LB media containing 50 μ g/ml ampicillin and grown at 37°C overnight. The culture was then diluted by half with fresh LB/ampicillin and cooled to room temperature for 1 hr, before inducing with 0.4mM IPTG for 6 h at 25°C.

Cells were recovered and incubated with 0.5 mg/ml lysozyme in 50 mM Tris (pH 8.0), 150 mM NaCl, 1% Tween-20, 0.1% 2-mercaptoethanol, 5 mM EDTA, 1 mM PMSF and a mixture

of other protease inhibitors obtained from Boehringer Mannheim (1697498) at room temperature for 0.5 h, followed by sonication. Cellular debris were pelleted by centrifugation at 27,500g for 10 min and the resulting supernatants were incubated with 30 ml of glutathione-Sepharose (Pharmacia) at 4°C overnight. The resin was then washed with 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, and 0.1% 2-mercaptoethanol until the OD 280nm reached <0.01. For removal of GST, the resin with immobilized GST-fusion protein was incubated with 10U of thrombin (Boehringer, Inc.) at 4°C in 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, 0.1% 2-Mercaptoethanol, and 2.5 mM CaCl₂ overnight. Released proteins were then purified on Mono Q (HR10/10, Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0 and dialyzed into chaperone assay buffer.

The ability of BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) to bind Hsc70/ATPase in solution was then examined. GST control or GST-BAG proteins were immobilized on glutathione-Sepharose and tested for binding to ³⁵S-labeled *in vitro* translated (IVT) proteins. Immunoprecipitation and *in vitro* GST-protein binding assays were performed as described by Takayama et al., *supra* (1997), using pCI-Neo flag or pcDNA3-HA into which human Bag-2 (clone #11) or human BAG-3 (clone #28) had been subcloned for *in vitro* translation of ³⁵S-L-methionine labeled proteins or expression in 293T cells. As shown in Figure 11B, ³⁵S-Hsc70/ATPase bound *in vitro* to GST-BAG-1, GST-BAG-2, and GST-BAG-3 but not to GST-BAG-1(ΔC) or several other control proteins. BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) also exhibited little or no binding to themselves or to each other, demonstrating that these proteins do not strongly homo- or hetero-dimerize or oligomerize. It should be noted, however, that BAG-2 (SEQ

ID NO:4) displayed weak interactions with itself in binding assays and produced a positive result in yeast two-hybrid experiments, demonstrating that it can have the ability to self-associate.

5 B. Binding of BAG proteins to Hsc70 in vivo

The ability of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins to interact in cells with Hsc70 was tested by expressing these proteins with N-terminal Flag epitope tags in 293T human epithelial cells using co-10 immunoprecipitation assays as described previously (Takayama et al., *supra* (1997)). cDNAs encoding the λ -phage cloned regions of BAG-2 and BAG-3 were subcloned in-frame into pcDNA3-Flag. Anti-Flag immune complexes prepared from 293T cells after transfection with plasmids 15 encoding Flag-BAG-1, Flag-BAG-2, or Flag-BAG-3 were analyzed by SDS-PAGE/immunoblot assay. As shown in Figure 10C, antiserum specific to Hsc70 detected the presence of BAG proteins associated with Hsc70, whereas control immune-complexes prepared with IgG1 as well as anti-Flag immune 20 complexes prepared from cells transfected with Flag-tagged control proteins, Daxx and Apaf-1, did not contain Hsc70 associated protein. These results further demonstrate that BAG-family proteins specifically bind to Hsc70.

25 C. BIACore assay of BAG protein binding to the ATPase domain of Hsc70

BAG-1 (beginning at residue 116 of SEQ ID NO:2) is known to bind tightly to the ATPase domain of Hsc70 (Stuart et al., *J. Biol. Chem.*, In Press (1998)). BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins were 30 therefore, examined for their ability to bind to Hsc70/ATPase. The affinity and binding kinetics of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) to Hsc70/ATPase was also compared to that of BAG-1 (beginning at residue 116 of

SEQ ID NO:2) for Hsc70/ATPase, using a surface plasmon resonance technique (BIAcore) which has been described previously (Stuart et al., *supra*, (1998) which is incorporated herein by reference).

5 BAG-family proteins were produced in bacteria and purified to near homogeneity as shown in Figure 12A and described above in Example I. The purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), -2 (SEQ ID NO:4), and -3 (SEQ ID NO:6) proteins were then immobilized
10 on biosensor chips and tested for their interactions with Hsc70 in the soluble phase. Kinetic measurements were performed using a BIAcore-II instrument with CM5 sensor chip and Amine Coupling Kit (Pharmacia Biosensor AB, Sweden). Briefly, for immobilization of proteins, the
15 sensor chip was equilibrated with HK buffer (10 mM Hepes (pH 7.4), 150 mM KCL) at 5 μ l/min, then activated by injecting 17 μ l of 0.2M N-ethyl-N'-(3-diethylaminopropyl)-carbodiimide and 0.05M N-hydroxysuccinimide (NHS/EDC) followed by 35 μ l of the protein of interest, in 10 mM acetate, pH 3.5-4.5. Excess NHS-ester on the surface was
20 deactivated with 17 μ l 1M ethanolamine-HCL (pH8.5). After immobilization, 5 μ l of regeneration buffer (50 mM phosphate (pH 6.8) and 4M GuHCl) was injected. For binding assays, Hsp70 (Sigma, H8778) was dissolved in HK buffer, and
25 injected at 10 μ l/min across the prepared surface at various concentrations. The surface was regenerated after each injection with 5 μ l of regeneration buffer. The rate constants k_{ass} and k_{diss} were generated with BIAevaluation softward 3.01 (Pharmacia Biosensor AB). Addition of Hsc70
30 to chips containing BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) resulted in concentration-dependent binding, as reflected by an increase in the Response Units (RU) measured at the chip surface (shown in Figure 3B). In contrast, Hsc70
35 failed to display interactions in BIAcore assays with a variety of control proteins as well as a mutant of BAG-1

lacking a C-terminal portion of the BAG domain which is required for Hsc70-binding (Figure 3B). Furthermore, flowing of various control proteins such as GST, BSA and Bcl-XL over the BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) chips resulted in negligible interaction. These results further demonstrate the specificity with which BAG-family proteins interact with and bind to Hsc70.

The rates of Hsc70 binding to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) proteins were similar, following pseudo first-order kinetics with estimated association rate constants (κ_a) of 2.1, 2.1 and $2.4 \times 10^5 \text{ M}^{-1} \text{ sec}^{-1}$, respectively. After allowing binding of Hsc70 to immobilized BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to reach plateau levels, the chaperone was removed from the flow solution and the dissociation rate was monitored. BAG-1 (beginning at residue 116 at SEQ ID NO:2) and BAG-2 (SEQ ID NO:4) exhibited similar dissociation rates, with relatively slow loss of Hsc70 from the chip surface, resulting in estimated dissociation rate constants (κ_d) of 3.0 and $5.0 \times 10^{-4} \text{ sec}^{-1}$, respectively (see Figure 3B). In contrast, Hsc70 dissociated more rapidly from biosensor chips containing BAG-3 (see Figure 3B), yielding an estimated κ_d of $1.7 \times 10^{-3} \text{ sec}^{-1}$. From the kinetic data, the apparent affinities (K_D = κ_d/κ_a) were calculated for binding of Hsc70 to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) and were estimated to equal about $K_D = 1.4 \text{nM}$, $K_D = 2.4 \text{nM}$, and $K_D = 7.4 \text{nM}$, respectively. These results demonstrate that the interactions of BAG-family proteins with Hsc70 occur with apparent affinities sufficient for physiological relevance.

EXAMPLE III

BAG-family proteins inhibit
Hsp70/Hsc70-dependent protein folding

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins inhibit Hsp70/Hsc70-dependent refolding of denatured proteins similarly to a BAG-1 (beginning at residue 116 of SEQ ID NO:2) protein.

The effects of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) protein on Hsp70/Hsc70-dependent protein refolding was determined using *in vitro* protein refolding assays similar to those described previously by Takayama et al., *supra*, 1998; Terada et al., *J Cell Biol.*, 139:1089-1095, 1997, which are incorporated herein by reference. Briefly, luciferase (20 μ M) was denatured in 25 mM Hepes-KOH, pH 7.2, 15 50 mM potassium acetate, 5 mM DTT, 6M guanidine hydrochloride at ~25°C for 1 h. Denatured luciferase was diluted 1:40 into 25 mM Hepes-KOH, pH 7.2, 50 mM potassium acetate, 5 mM DTT. Hsc70 (1.8 μ M), DnaJ (StressGen, Inc.) (0.9 μ M), and various purified recombinant proteins as 20 indicated were added to refolding buffer (30 mM Hepes-KOH, pH 7.6, 120 mM potassium acetate, 3mM magnesium acetate, 2 mM DTT, 2.5 mM ATP) with 0.2 volume of diluted denatured luciferase to a final concentration of 0.1 μ M. Luciferase activity was measured after 1.5 hr incubation at 35°C.

25 The combination of Hsc70 and DnaJ resulted in ATP-dependent refolding of chemically denatured firefly luciferase, with function of over half the denatured enzyme restored in a 90 minute reaction, as monitored by a chemiluminescence assay. In contrast, neither Hsc70 nor 30 DnaJ alone were able to induce substantial refolding of denatured luciferase. Furthermore, little spontaneous

restoration of luciferase activity was observed with control proteins, BSA, GST or Bcl-XL (see Figure 4A).

Addition of recombinant purified BAG-1 (beginning 5 at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to the above assays in amounts equimolar to Hsc70 (1.8 μ M) resulted in striking inhibition of luciferase refolding. BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) displayed somewhat greater inhibitory 10 activity than BAG-1 (beginning at residue 116 of SEQ ID NO:2) as shown in Figure 4A. In contrast, the BAG-1 (Δ C) protein, which fails to bind Hsc70 as well as several other control proteins, had no effect on luciferase refolding.

In an additional refolding assay, described 15 previously by Minami et al., J Biol. Chem. 271:19617-24, 1996), purified Hsc70 and human DnaJ homolog Hdj-1 (Hsp 40) were used with additional cofactors provided in reticulocyte lysates (5% v:v) to produce a system capable of refolding denatured luciferase. Briefly, additional 20 cofactors included, recombinant Luciferase (Promega: QuantiLum TM), that had been heat denatured at 42°C for 10 min, 1.8 μ M Hsc70 (Sigma; purified from bovine brain), 0.9 μ M Hsp40, and various recombinant purified proteins. Luciferase activity was measured (Promega luciferase assay 25 kit) using a luminometer (EG&G Berthold, MicroLumat luminometer, Model #LB96P). All results were normalized relative to non-denatured luciferase that had been subjected to the same conditions. Control reactions lacking ATP, Hsc70, or Hsp40 resulted in negligible 30 luciferase refolding.

Various amounts of purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6), relative to amounts of Hsc70 were used in the above-described protein refolding assay. Addition of 35 BAG-family proteins resulted in a concentration-dependent

inhibition of Hsc70 chaperone activity. Furthermore, the BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) inhibition of Hsc70 chaperone activity was demonstrated to be as potent as that observed for BAG-1 (beginning at residue 116 of SEQ 5 ID NO:2). In contrast, the BAG-1 (Δ C) mutant as well as other control proteins did not suppress Hsc70-mediated refolding of denatured luciferase. These results indicate that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) can inhibit Hsc70/Hsp70 dependent protein refolding activity to 10 the same extent as BAG-1 (beginning at residue 116 of SEQ ID NO:2).

B. BAG competes with Hip for binding to Hsc70.

It is known that BAG-1 competes with Hip for binding to Hsc70, with these proteins exerting opposite 15 effects on Hsc70-mediated protein refolding (Hohfeld, J., and Jentsch, S., *Embo J.*, 16:6209-6216, 1997, which is incorporated herein by reference). In order to determine whether BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) also compete with Hip for binding to Hsc70, refolding assays 20 were performed as described above in the presence of Hip protein.

Hip was purified as His₆-protein. The fusion protein was induced from pET28-Hip (V. Prapapanich et al., *Mol Cell Biol.*, 18:944-952, 1998, which is incorporated 25 herein by reference) with 0.1 mM IPTG at 25°C for 6h in BL21 cells. Cells from 1L of culture were resuspended into 50 ml of 50 mM Phosphate buffer (pH 6.8), 150 mM NaCl, and 1% (v/v) Tween-20 and then incubated with 0.5 mg/ml lysozyme at 25°C for 0.5h, followed by sonication. After 30 centrifugation at 27,500g for 10 min, the resulting supernatant was mixed with 15 ml nickel resin (Qiagen, Inc.) at 4°C for 3 h with 25 mM imidazol. The resin was then washed with 50 mM phosphate buffer (pH 6.8), 25 mM imidazol, 150 mM NaCl and 0.1% Tween-20 until the OD_{280nm}

reached a value of <0.01. His₆-Hip protein was eluted with 250 mM imidazol in washing buffer (Qiagene, Inc.) and purified on Mono Q (HR10/10 Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0, followed by 5 dialysis in chaperone assay buffer.

In the refolding assay reactions, addition of purified Hip at equimolar concentrations relative to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) (1.8 μ M) completely negated 10 the inhibitory effects of the BAG-family proteins on refolding of denatured luciferase (see Figure 4C). These results demonstrate that the suppression of Hsc70 chaperone activity by BAG-family proteins is reversible, and that Hip antagonizes the effects of not only BAG-1 (beginning at 15 residue 116 of SEQ ID NO:2), but also of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6).

In summary, these results demonstrate that BAG-family proteins all contain a conserved BAG domain near their C-terminus that binds Hsc70/Hsp70, and that human 20 BAG-family proteins can bind with high affinity to the ATPase domain of Hsc70 and inhibit its chaperone activity through a Hip-repressable mechanism.

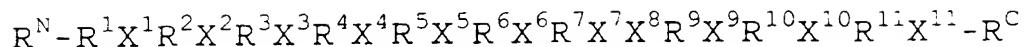
EXAMPLE IV

EXPANDED NUCLEIC ACID AND AMINO ACID SEQUENCES
25 FOR HUMAN BAG-3, BAG-4 AND BAG-5

Following the procedures disclosed herein, the nucleic acid and amino acids sequences to human BAG-3, BAG-4 and BAG-5 were further expanded. The expanded sequences for BAG-3, BAG-4 and BAG-5 are shown in 30 Figures 15, 16 and 17, respectively, with their respective sequence identification numbers, "SEQ ID NO"s.

We claim:

1. A compound of the formula,



wherein,

5 R^N is a group of about 1 to 552 independently selected amino acids;

10 R^1 is a group of 3 independently selected amino acids;

15 X^1 is an amino acid with a charged or uncharged R group;

20 R^2 is a group of 7 independently selected amino acids;

25 X^2 is an amino acid with a charged R group;

30 R^3 is a group of 5 independently selected amino acids;

X^3 is an amino acid with an apolar R group;

R^4 is a group of 3 independently selected amino acids;

X^4 is an amino acid with charged R group;

R^5 is a single independently selected amino acid;

X^5 is an amino acid with apolar or uncharged R group;

R^6 is a group of 15 independently selected amino acids;

X^6 is an amino acid with a charged or uncharged R group;

R^7 is a group of 2 independently selected amino acids;

X^7 is an amino acid with a charged R group;

X^8 is an amino acid with a charged R group;

R^9 is a group of 2 independently selected amino acids;

X^9 is an amino acid with an apolar R group;

R^{10} is a group of 3 independently selected amino acids;

X^{10} is an amino acid with an uncharged R group;

5 R^{11} is a group of 2 independently selected amino acids;

X^{11} is an amino acid with an apolar R group; and

R^C is a group of about 1 to 100 independently selected amino acids.

2. A substantially purified nucleic acid
10 molecule having a nucleotide sequence corresponding to or complementary to at least 20 nucleotides from a nucleotide sequence selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).

15 3. The nucleic acid of claim 2 having a nucleotide sequence corresponding to or complementary to a nucleotide sequence that encodes a functionally active BAG family protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ
20 ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).

25 4. The nucleic acid of claim 3 selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).

30 5. The nucleic acid of claim 3 complementary to a nucleotide sequence that encodes a functionally active BAG protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).

6. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:3).

7. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:5).

8. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:7).

5

9. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:9).

10. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:19).

10 11. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:21).

12. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:23).

13. A substantially purified BAG family protein
15 encoded by the nucleic acid molecule of claim 1.

14. A substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ 20 ID NO:22) and (SEQ ID NO:24) or a fragment, a derivative or a mimetic thereof.

15. A substantially purified protein corresponding to the amino acid sequence of 157 to 204 of (SEQ ID NO:2).

25 16. A substantially purified protein corresponding to the amino acid sequence of 272 to 319 of (SEQ ID NO:2).

17. A substantially purified protein corresponding to the amino acid sequence of 164 to 211 of (SEQ ID NO:4).

18. A substantially purified protein 5 corresponding to the amino acid sequence of 418 to 510 of (SEQ ID NO:20).

19. A substantially purified protein corresponding to the amino acid sequence of 378 to 457 of (SEQ ID NO:22).

10 20. A substantially purified protein corresponding to the amino acid sequence of 6 to 97 of (SEQ ID NO:24).

15 21. A substantially purified protein corresponding to the amino acid sequence of 180 to 257 of (SEQ ID NO:24).

22. A substantially purified protein corresponding to the amino acid sequence of 272 to 349 of (SEQ ID NO:24).

20 23. A substantially purified protein corresponding to the amino acid sequence of 362 to 444 of (SEQ ID NO:24).

25 24. A pharmaceutical composition comprising a nucleic acid molecule of claim 1 useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.

25. A method of modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function by administering a nucleic acid molecule of claim 1.

26. A pharmaceutical composition comprising a substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), 5 (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24), or a fragment, a derivative or a mimetic thereof, useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.

10 27. A method of modulating tumor cell proliferation by administering a pharmaceutical composition of claim 26.

15 28. A method of modulating cell migration and metastasis by administering a pharmaceutical composition of claim 26.

29. A method of modulating steroid hormone receptor function by administering a pharmaceutical composition of claim 26.

30. A substantially purified antibody that 20 specifically binds to a BAG family protein of claim 14.

31. The antibody of claim 30, wherein said antibody is a monoclonal antibody.

32. A method for detecting the presence of a BAG family protein in a sample, comprising the steps of:

- a. obtaining the sample;
- b. adding to said an antibody of claim 11 under suitable conditions for the binding of said antibody with the BAG family protein; and
- c. detecting said bound BAG family protein.

10 33. A method for detecting the presence of a first nucleic acid molecule that encodes a BAG family protein in a sample, comprising the steps of:

- a. obtaining the sample;
- b. adding to said sample a second nucleic acid molecule capable of hybridizing with said first nucleic acid molecule under suitable conditions for the binding of said second nucleic acid molecule with said first nucleic acid molecule; and
- c. detecting said hybridized first and second nucleic acid molecules.

20 25 34. A method of determining the risk of metastatic spread of cancer or prognosis of cancer patients by determining the level of expression of a BAG-family protein.

FIGURE 1

FIGURE 2A

GCGGCCGCG TGTGGGAG TCTCCCGG TTGCCCGG GCGGCCGCG GCGGCCGCG GCGGCCGCG GCGGCCGCG 90
 CAGGCCGCG TCCACTGGCT GCGGCCGCG GCGGCCGCG CTCCTGGCTA CCCCCGGTCG GAGGCTTAGA TGGCTAGGC GAGATCAAC
 A K A N E G R F C R S S M A D R S S A L L E S L D Q L E L 180
 CCTAAGCCCA ACCAGGGCG CTTCTGCCG TCCCTCTCCA TGGCTGACCG CTCAGGCCG CTGCTGGAGA GCCTGGACCA GCTGGAGCTC 270
 AGCGTTGAG CTTTGAGGAG AGCGGCGACT GCTTGAGG AGAGGAGAGA AATCTTCTG GAAATGATTC AGAGTATCCA AATATGCCAG
 A V E A L R E A A T A U E Q E K E I L L E M I H S I Q H S Q 360
 GCGATGGCC AGATCAGTGA CGGAGGAGA GAGGATTAAT ATCTGACTGC AAGCCGTTG ATGGGAGAGA CTCTCACCGT TGAATGTCGA
 D H R Q I S D G E R E E L H L T A N R L H G R T L T U E V S 450
 GTAGGAAACAA TTAGGAAACCC CGAGGAGGAGA GAAATGCCAA AGCATGCCAC AGGGATTTATT GATGAGGGTGG TCAATAGTT TCTGGATGAT
 V E T I A N P Q Q E S L K H A T A I D E U V N K F L D D 540
 TTGGGAAATG CCAAGAGTCA TTAAATGTCG CTCTACAGTCA ATGTTCATC TGAGGTGCCA CATGGGCCAG TTGATCAGAA GTTTCATCC
 L G H A K S H L H S L Y S A C S S E U P H G P U D Q K F Q S 630
 ATAGTAAATTG GCTGTGCTCT TGAAGATCAG AGAGGAGATT AGAGGAGATT AGAGGAGATT AGAGGAGATT AGAGGAGATT AGAGGAGATT
 I U I G C A L E D Q K K I K R R L E T L R N I E H S D K A 720
 ATCAGGCTAT TAGGCTATTC TGAAGAGCT GGTCCAAAATCTGGCAAGA AATGCTGAA AGCGGATTCA ATTAGTCTTC AAGCCTAAGAA 810
 I K L L E H S K G A G S K T L Q Q N A E S R F N

FIGURE 2B

900
CTGATAGTTG TTTCAGATGA CGGAAATTATT CCGATCAGATA TCTTCAGTTT TGTGATATAC AAGACTAGCA ATATTTTAT TATCTATCTA 990
GAGATTTTTT AGATTGATT CTTGTCCTGT ACTAGGATCT AGCATATTTC ACTATTCTGT AGATGATAC ATAGTTTGTG CGGAAACAA 1080
ACGTTTACGCT AGGGGCAAA AGGATGACTG CTTTTTCTG TCTGGATGG ATCACGGAG TCACCTTGG CATTAGTTT ACTGAAATT 1170
CTTTTACTGG 1179

GCATTACAC AATACACACG GTGTAAATAT GATAAAATAC TATTTTATT GATAACTAGT TCTTTGTAGT GATAAACCAC TTAGTTGACA

FIGURE 3

GCGAGCTCC GCGTGCAGCC CGCGGGGGGG CGCGCTCTCT CGCGACTGAG CGAGAGTTT CTAGGGGGGC AGTGGCTACC TCCCTTATC
 R E L R I Q P R R A A M F S G L O Q K F L A G Q L L P P F I 90
 TCCCTCTTCC CCTCTGGCGAG CGACGAGCT ATTTCAGACG ACTTCCACCC CTCTCTGGCC AGCTCACCCCG CGCTTATATCT TCTTAAAGGT
 S S F P S G S E E R I S R H F H P S L A T S P P P L I H K G 180
 CGCGGGGGGG CGCTTCCCGG AGCGTCCGG CGCGGAGGG CGCGCACCGC CGCGGGGGGG CGAGAGACTC CGCGGGGGGG CGAGAGGCC
 R R R R L P G H U G G G E G P T R A R R P E T R R P E P A P 270
 CGACGGGGCG CGCGGGGGGG CGACGCGCA CGACGCGATCA CGCGGGGGCG CGCTCGGGCC ATGATCCAGG TGCGTCCGG CGACGTCAC
 R T R A P R G R P Q P S M S A R T H S P M H Q V A S G H G D 360
 CGCGACCTT TGCCCCCGG AGAGTCCACG CGAGAGCGG CTGGCCCTTC TTCTGGGCC AGACAGCGG CGACGACTACG
 R O P L P P G H E I K I D P Q T G H P F F V D H N S R T T T 450
 TGGACACCC CGCGGCTGCC CTCTGAGGGC CGCAAGGAGA CTCCATCTC TGCGATGGC CCTTCCCCGG AGCGCTCTAG CCTGGCCCGT
 W N D P R U P S E G P K E T P S S A N G P S R E G S R L P P 540
 GCTAGGGAG CGACCCCTGT GTACCCCCAG CTCCGACCGAG CGTACATTC CATTCCYTG CTCCATGGAG CGCTGAGAA CGGGAGGTG
 R A E G H P U Y P Q L R P G Y I P I P U L H E G A R E N R Q U 630
 CGCCCTTTCG ATGCTCTATCC CGCGCTGGG AGCGAGCGAT TCGACGACTCA CGCGGAGCA CGCGCTCCG AGAGGTCCCA GTCACCTCTG
 H P F H U Y P Q P G M Q R F R T E A R A A R A R P Q R S Q S P L 720
 CGCGGCGATCG CGACGACAC CGACGAGAT AACAGTGTG CGAGGTGGC AGCGGGGGCG CGACGCGCTC CGCGGACCT
 R O M P E T T Q P O K Q C G Q U A A R A A R A Q P P A S H G P 810
 GAGCGGCTCC AGTCTCCACG TGCTCTACG TGCTCTCTC CATCTCTC CGCGAGCTG CCTTCTCG CGAGGAGG CCTGGCCAGT
 E R S Q S P A R S D C S S S S S A S L P S S G R S S L G S 900
 CACCAAGCTCC CGCGGGGGTA CATCTCTATT CGCGTACATAC ACCAGGAGA CGTACCGGG CGACGAGGCC AGCGCTCTT CGACGAGCC
 H Q L P R G Y I S I P U H E Q N U T R P A A Q P S F H K A 990
 CAGAAGACGC ACTACCCAGC CGAGAGGGT CGTACCCAGA CGACGAGGC TGTTACCGG CGACGAGGCC AGAGTGGCTG CGAGGCCGG
 Q K T H Y P A Q R G E Y Q T H Q P U Y H K I Q G D D H E P R 1080
 CGCCCTGGGG CGCGATCCCC GTTCAGGTCA TCTCTCCAGG TGCGATGGC CGGGGGGGC TCACCGAGCA CGACGAGCAC CGACCTCCAC
 P L R A A S P F R S S U Q G A S S R E G S P A R S S T P L H 1170
 TCCCGCTCCG CCTCCGTGT CGACACCGTG TGCGACAGGC CTGAGCGCC CATGACCAT CGAGGAGCTG CACCTGTTTC CGAGCTGAA
 S P S P I A V H T U U D R P Q Q P M T H R E T A P U S Q P E 1260
 AACAAACCG AGAGTAAACG AGGCCCTGTT CGACGAGAC CGCCCTCTGG ACACATCCCA ATTCAAGTGA TCCCGAAAGA CGTGGATTCT
 H K P E S K P G P U G P E L P P G H I P I Q V I R K E U D S 1350
 AACACCTGTT CGACGAGCC CGCGCTCCC TCTGAGGAGG TAGAGGTGA AGTTCGGCGT GCTCCAGTTC CTGGCTCTCC CGCGGGCT
 K P U S Q K P P P P S E K U E U K U P P A P U P C P P P S P 1440

CGCCCTCTG CTGTCGGCTC TTCCCCAG AGTGTGGCTA CGAGAGAG CGCGGGGGCG AGCGCTGCC CGCGAGAGC TACACCTGAA
 G P S A U P S S P K S U A T E E R A P S T A P R E A T P P 1530
 AACCCAGAG AGCCCGAGGC CGCCCGAGA CATCGAGAG TCGTACGAT CGACGAGCTC CTGGAGAGG TGAGGGCT CGACGAGGT
 K P G E A E R P P K H P G U L K V E A I L E K V Q G L E Q A 1620
 GTAGACACT TTGAGGGCA GAAGACTGAG AAGAAGTACG TGATGATCGA AGAGTATTTG ACCAAGAGGC TGCTGGCCCT CGATTGAGTG
 V D H F E G K K T O K K Y L H I E E Y L T K E L L A L D S V 1710
 CGACCGAGC CGACGACGCC TGCGCTAG CGCGGGAGAG AGCGTGTAG CGACGAGCTG ACCATCTCTG AACACATCA AGACGAGGCC
 O P E G R A D V R Q A R R D G U R K U Q T I L E K L E Q K A 1800
 ATTGTGTCC CGGGCTCATG CGACGAGCTG CGACGAGCT CGACGAGAT CGACGAGCTG AGCGCATCAT CGAGGATGGT
 I D U P G Q U Q U Y E L Q P S H L E A D Q P L Q A I H E M G 1890
 CGCGTGGAG CGACGAGAG CGACGAGATG CGACGAGATC CGACGAGAG CGACGAGAG CGACGAGGCC
 A U A A O K G K K M A E D P H T E T Q Q P E A T A A A 1980
 ACTTCAACG CGACGAGCT CGACGAGCC CGCGTGTAG CGACGAGCT CGACGAGCTG CGCGTGTAGA CGACGAGCTG CGACGAGCT
 T S M P S S M T D T P G N P A P 2070
 GTCTTTAGG CATTTTAGT CGATGATTT CGACGACTT AGGTGAGTC GTTTCGTTA CGCTGCTTGGT AGCGACTACT CGGGTGGCC
 AACACATATA AGCGCTTAA AGCGAGATG ATGCTTTCT CGATGATTT TACTCTGGTA CGATGATATA AGCGATTTGT GTTTGGAGA
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2160
 AACACATATA AGCGCTTAA AGCGAGATG ATGCTTTCT CGATGATTT TACTCTGGTA CGATGATATA AGCGATTTGT GTTTGGAGA
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2250
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2340
 AACACATATA AGCGCTTAA AGCGAGATG ATGCTTTCT CGATGATTT TACTCTGGTA CGATGATATA AGCGATTTGT GTTTGGAGA
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2430
 AACACATATA AGCGCTTAA AGCGAGATG ATGCTTTCT CGATGATTT TACTCTGGTA CGATGATATA AGCGATTTGT GTTTGGAGA
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2520
 AACACATATA AGCGCTTAA AGCGAGATG ATGCTTTCT CGATGATTT TACTCTGGTA CGATGATATA AGCGATTTGT GTTTGGAGA
 GTTTAACCGG GTGGCTTGGT CGACGAGCT CGACGAGCT CGACGAGAT CGACGAGCTG CGACGAGCTT TTGAGCTCT
 CGACGAGCC CGACGAGCT CGACGAGAT CGACGAGAT AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG
 TTGATGATATA AGATGAGGAG AGATGAGGAG AGCGATGAGG AGATGAGGAG AGATGAGGAG AGATGAGGAG AGATGAGGAG
 AGATGAGGAG 2528

ACGATATCCT	CTAAGACCAA	GATTGCAAA	GGCAAGGAA	GGGAGTTTT	GGATTCCTTAT	ACGATGGAG	CCTATGGCC	ACGATACCCC	CCAGCCCTG	90
GGCGGAAATAC	TCCTCTATAC	TGAGGGGCTT	ATTATGCAAC	TGGTTAACT	CAGACCAAGTT	ACTCCACAGA	AGTCCACAGT	ACTTACCGTT		180
CATCTGGCAA	CAGCCCAACT	CGAGCTCTTC	GTGGGGTCTA	TCCCAGGAG	GAATGCAAG	ACTGTCAAG	ACTGAGGCAC	CCCGCTTTAA	GGGCAAGTT	270
CGAGGATATC	GGCCTTCACA	GAACCTGGAA	ATGACCTGCA	CCCATTTATCC	TTATGGAGAT	GGTATCGTA	GTGTTCCACAA	ATCGAGGCCG		360
					H E M	V I U	U F H	N H B	A	
L Y D	H K K	D A H	A S P	G M A	Y G M	G G R	Y P H	P S S	A P S A	450
P P G	H L Y	H T E	S T S P	W P S S	G S P	Q S P	P S P	P V	U Q Q	540
GGCGGAGAT	TCTTCATACC	CCATAGGCCA	ATCGATCAAA	AGCATGAAAC	GGCGACACTT	TCCCTTCAGT	GTCCATCGT	ACGATCCTC		630
P K D	S S Y	P Y S	Q S D	Q S M	N R H	N F P	C S U	H Q Y	E S S	
G T U	N N D	D S D	L L D	S Q U	Q V Y	S A E	P Q L	Y G N A	T S D	720
CCATCCAC	ATTCACACATC	AAAGTAGGCG	TCTTCCTGAA	GGATGTAC	CTTCAGATCA	AGTACTCTCT	GCCTCGCTG	TATGGTATG	CCACCAATGAA	
H P H	N Q D	Q D Q	S S S	L P E	E C V P	S D E	S T P	P S I	K K I I	810
ACATGTGCTG	GAGGAGGTCC	AGTATCTTA	ACGAGAGTA	GGAGATTTC	TAGGAAAGAA	GGACGACAAA	GCATACGGC	TTCTGAAAC		900
H U L	E K V	Q Y L E	Q E U	E E F U	G K K T	D K A	Y W L	L E E		
AAATGCTTAC	AAGGGAACTTT	GGGAACTGAA	TTCAATTGAA	ACTGGGGGCC	AGGACTCTGT	ACGGCGGCC	AGGAGAGGG	CTGTTTGTAA		990
H L T	K E L L	E L O	S U E	T G G Q	D S U	R Q A	R K E	A U C K		1010
I Q A	I L E									

GAGGAAATAAA AAGATGACCTT CTCCCAACAC AACACCCTTC TGAATTGTCAC CTGAGGCTCA AAACGCAATT GCAGGGTTA ATTGGACAT 90
 E I K N E L L Q A Q N P S E L Y L S S K T E L Q G L I G Q L
 TGGATGAGGT AAGTNTGAA AAAAACCCCT GCATCCGGAA AGCCAGGAA AGAGGAGTGA TCGAGGTGCA AACTCTGTC ACATATATTG 180
 D E U S X E K N P C I R E A R R R A U I E V Q T L I T Y I D
 ACTTGAACCA GGCCTTGAG AAAAGGAAAGC TGTGTTGCTG TGAGGACAC CCATCCCATA AAGCCCTCTG GAAAGTCCCTT GGAACATTG 270
 L K E A L E K A K L F A C E E H P S H K A U W N U L G N L S
 CTGAGATCCA GGGAGGAGTT CTTTCATTG ATGGAAATCG ACGGATAG ARCTCATAG GGCTGGAGA CCTGCTCACCC AAGGAGCTGC 360
 E I Q G E U L S F D B N R T D K N Y I R L E E L L T K Q L L
 TAGCCCTGG A TGCTGTTGAT CGGAGGGAG AGAGGAGAGTGA TGGGCTGCC AGGAAACAGC CTGGAGGGCT TGGCGAGGAT ATTCTCTAGCT 450
 A L D A U D P Q G E E K C K A A R K Q A U R L A Q N I L S Y
 ATCTGACCT GAAATCTGAT GAAATGGAGT ACTGAAATAC CAGAGATCTC ACTTTGATA CTGTTGGCA CTTCATATGT GCTTCTATGT 540
 L D L K S D E W E Y
 ATACGAGCT TTCAAGTTCAT TGATTTATAC GTGCAATTATTT CAGTCTCAGT ATTTATGATT GAGGCAATT CTATTCAAGTA TCTGCTGCTT 630
 TTGATGTTGC AAGACAAATA TCATTAACGC ACGTAACTT TTCCATTGG ATCAAAAGA 689

FIGURE 6A

ATGTCTTCGGCTCTCGTTGAAATATTCACCTTCTTTCCAGCTTTCCCCATCTGACCT
GCTTTGGTTTT
CGAGAAAACCACGTTCAAATCAGCGACATCTCTCAAATTGAGATCATAGGCTTTGAAGATTG
CTCAAATTATG
CTTCTCATATTGCATGAGCATTGTGAAGCCCGGTCAACCAAAGCATTTCACCCATCA
CAATGATTAT CATTTCCTTAAAATT

8 / 35
FIGURE 6B

MKVNVSCSSV	QTTIDILEEN	QGEDESILTL	GQLRDRIATD	NDVDVETMKL	50
LHRGKFQGA	DDVSLSTLNF	KENDKIIIVMG	GKNALVDDAG	FKMLMQYEKH	100
NLSNLQKAYD	LNLRDVADLE	RGFLEPKQV	EMGKKLEKKV	KYFNEEAERH	150
LETLDGMNI I	TETTPENQAK	RNREKRKTLV	NGIQTLLNQN	DALLRRLQEY	200
QSVLNGDIPE					210

FIGURE 7A

ATGCCAGTCG	TGAACATACC	AATCAAAATA	CTTGGTCAGA	ATCAATCACA	50
TAGTCGAAGT	AACTCCTCGT	CTTCTGTTGA	CAACGATCGA	AATCAACCCAC	100
CACAGCAGGCC	ACCTCAACCG	CAACCACAAAC	AGCAATCTCA	GCAACAATAC	150
CAGCAGGCTC	CAAACGTGAA	TACCAATATG	CATCATTCGA	ACGGATTCTC	200
ACCTTAACCTTC	CCATCTCGTA	GTCCTATTCC	GGACTTTCCC	AGTTTTTCAT	250
CTGGGTTCCC	AAACGATTCT	GAATGGCTT	CGAATTTC	GTCGTTCCA	300
AATTTCCTAA	GTGGATTCTC	AAATGGAAGT	TCTAATTTC	CTGATTTCC	350
AAGATTCGGA	AGAGATGGAG	GAATATGCC	AAACCCACCG	ATGCAAGGAT	400
ACAGGAGAAG	TCCAACACCA	ACATCAACTC	AATCTCCAAC	TTCTACATTA	450
AGACGCAACT	CTCAGCAGAA	TCAAGCTCCT	CCACAATATT	CTCAGCAACA	500
ACCACAAACAA	GCTCAACAAAC	GTCAGACAAAC	TCCTCCGTCA	ACAAAAGCTT	550
CATCTCGACC	ACCATCTCGT	ACTCGTGAAC	CAAAGGAACC	TGAGGTACCC	600
GAGAGACCGAG	CAGTTATTCC	ATTGCCATAT	GAGAAGAAGG	AGAAAACCACT	650
GGAGAAGAAA	GGTAGTCGTG	ATTCTGGAAA	GGGTGATGAG	AACCTGAAAG	700
AGAACATTTGC	CAAGATCACG	ATCGGAAAGA	ATAATTGCGA	GTTATGTCCG	750
GAACAAGAAA	CGGACGGCGA	CCCATCTCCA	CTAACCTCCC	CAATCACCAG	800
AGGAAAGCCA	AAGAGAGGAA	AGAAACTTCA	ACGTAATCAA	AGTGTGTTG	850
ATTTCATATGC	CAAGACAATT	GTTACTTTGG	ATAAAATTGA	ATTACAAGTT	900
GAGCAGTTGA	GAAAAAAAGC	TGCTGAACTC	GAAATGGAAA	AAGAGCAAAT	950
TCTTCGTTCT	CTAGGAGAAA	TCAGTGTCA	TAAC TGATG	TTCAAACGG	1000
AAGAATGTGA	TCGTGAAGAG	ATTGAAGCAA	TCACTGACCG	ATTGACAAAAA	1050
AGAACAAAGA	CAGTTCAAGT	TGTTGTCGAA	ACTCCACGAA	ATGAAGAACAA	1100
GAAAAAAAGCA	CTGGAAGATG	CAACTTTGAT	GATCGATGAA	GTCGGAGAAA	1150
TGATGCATTG	GAATATTGAA	AAGGCTAAGC	TGTGCCTACA	AACCTACATG	1200
AACGCCTGTT	CGTACGAAGA	AACTGCTGGA	GCCACCTGCC	AAAACCTCTT	1250
GAAGATCATA	ATTCACTGCG	CTGCTGATGA	TCAGAAACGC	ATCAAGCGTC	1300
GTCTGGAAAA	TCTGATGTCT	CAAATTGAGA	ATGCTGAGAG	AACGAAAGCA	1350
GATTTGATGG	ATGATCAAAG	CGAATAG			1377

FIGURE 7B

MPVVNIPIKI	LGQNQSHSRS	NSSSVNDNR	NQPPQQPPQP	QPQQQSQQQY	50
QQAPNVNTNM	HHSGNGFSPNF	PSRSPIPDFP	SFSSGFPNDS	EWSSNFPSPFP	100
NFPNGFSNNGS	SNFPDFPRFG	RDGGLSPNPP	MQGYRRSPTP	TSTQSPTSTL	150
RRNSQQNQAP	PQYSQQQPOQ	AQQRQTTPPS	TKASSRPPSR	TREPKEPEVP	200
ERPAVPLPY	EKKEKPLEKK	GSRDSGKGDE	NLEENIAKIT	IGNNCELCP	250
EQETDGDPS	LTSPITEGKP	KRGKKLQRNQ	SVVDFNAKTI	VTLDKIELQV	300
EQLRKAAEL	EMEKEQILRS	LGEISVHNCM	FKLEECDREE	.IEAITDRLTK	350
RTKTVQVVVE	TPRNEEQKKA	LEDATLMIDE	VGEMMHSNIE	KAKLCLOTYM	400
NACSYEETAG	ATCQNFLKII	IQCAADDQKR	IKRRLENLMS	QIENAERTKA	450
DLMDQSE					458

FIGURE 8A

ATGTCAGAAA	AGACTAGCAC	AGTTACAATA	CACTATGGAA	ATCAGCGATT	50
TCCGGTAGCA	GTCAATCTAA	ATGAGACGTT	AAGTGAAC TG	ATTGATGATT	100
TACITGAAAC	GACTGAGATT	TCTGAGAAGA	AAGTCAAGCT	TTTTTACGCT	150
GGCAAGCGTT	AAAAGACAA	AAAAGCCTCG	TTATCAAAAT	TGGGTTAAA	200
AAATCATAGT	AAAATTCTAT	GTATAAGACC	ACATAAGCAA	CAACGAGGTT	250
CCAAGGAAAA	AGACACGGTT	GAGCCCGCTC	CGAAAGCGGA	AGCGGAGAAT	300
CCTGTATTTT	CGCGTATTTT	TGGAGAAATA	AAAGCCATCG	ATCAGTATGT	350
TGACAAAGAA	CTTTCCCCCA	TGTACGACAA	TTACGTAAT	AAACCGTCGA	400
ACGATCCAAA	GCAGAAAAAC	AAACAGAAAC	TAATGATAAG	TGAAC TACTT	450
TTACAACAGC	TTTTAAAATT	GGATGGAGTT	GACGTACTGG	GCAGCGAGAA	500
ATTGCCTTTT	GAACCGGAAGC	AACTTGTTC	TAAGATCCAA	AAAATGTTGG	550
ATCACGTTGA	CCAAACAAAGC	CAAGAAGTGG	CCGCATAG		588

FIGURE 8B

MSEKTSTVTI	HYGNQRFPVA	VNLNETLSEL	IDDLLETTEI	SEKKVKLFYA	50
GKRLKDKKAS	LSKLGLKNHS	KILCIRPHKQ	QRGSKEKDTV	EPAPKAEAEN	100
PVFSRISGEI	KAIDQYVDKE	LSPMYDNYVN	KPSNDPKQKN	KQKLMISELL	150
LQQLLKLDGV	DVLGSEKLRF	ERKQLVSKIQ	KMLDHVDQTS	QEVAQ	195

FIGURE 9A

ATGTCTTTTT	TTACCCAGTT	GTGTTCTATG	GATAAAAAAT	ATTGGATCTC	50
TCTAGCTGTA	TTGTCAGTTA	CTGTTTGAT	TAGCGCATT	TTGAAAAAAGA	100
GAGCTACTGA	AACCGAAGAT	ATTGTCGTTG	TTCATTACGA	TGGCGAAAAG	150
TTGAATTGGT	TGTTGCGACA	ACCAAGGCTG	AATATGGTTT	CTTACACTAG	200
TTTTCTTCGT	CGCGTGTGCA	ACGCATTTC	AGTAATGCC	GACAAAGCGT	250
CTCTCAAGTT	AAACGGGGTG	ACCCCTCAAGG	ATGGTTCACT	TTCCGACCAA	300
AATGTGCAAA	ATGGAAGTGA	ATTAGAGCTC	GAATTACCCA	AACTGAGCCC	350
GGCAATGCAA	CAAATTGAAG	CATATATAGA	TGAGCTTCAA	CAGGATCTCG	400
TCCCTAAAAT	TGAAGCCTTC	TGCCAATCGT	CTCCCGCTTC	GGCACAAAGAT	450
GTTCAAGATT	TGCATACACG	CCTTAGTGAA	ACATTGTTGG	CTAGGATGAT	500
AAAATTAGAT	GCTGTTAATG	TTGAAGACGA	CCCAGAAGCT	CGTCTTAAAA	550
GAAAAGAAGC	TATTCGTTTA	TCTCAACAAAT	ATTTGAGTAA	ACTAGATTCC	600
ACCAAGAACATC	AAAACAAATG	A			621

FIGURE 9B

MSFFTQLCSM DKKYWISLAV LSVTVLISAL LKKRATEDTED IVVVHYDGEK	50
LNFVLRQPRRL NMVSYTSFLR RVCNAFSVMP DKASLKLNGV TLKDGSLSDQ	100
NVQNGSELEL ELPKLSPAMQ QIEAYIDELQ QDLVPKIEAF CQSSPASAQD	150
VQDLHTRLSE TLLARMIKLD AVNVEDDPEA RLKRKEAIRL SQQYLSKLDs	200
TKNQNK	206

FIGURE 10A

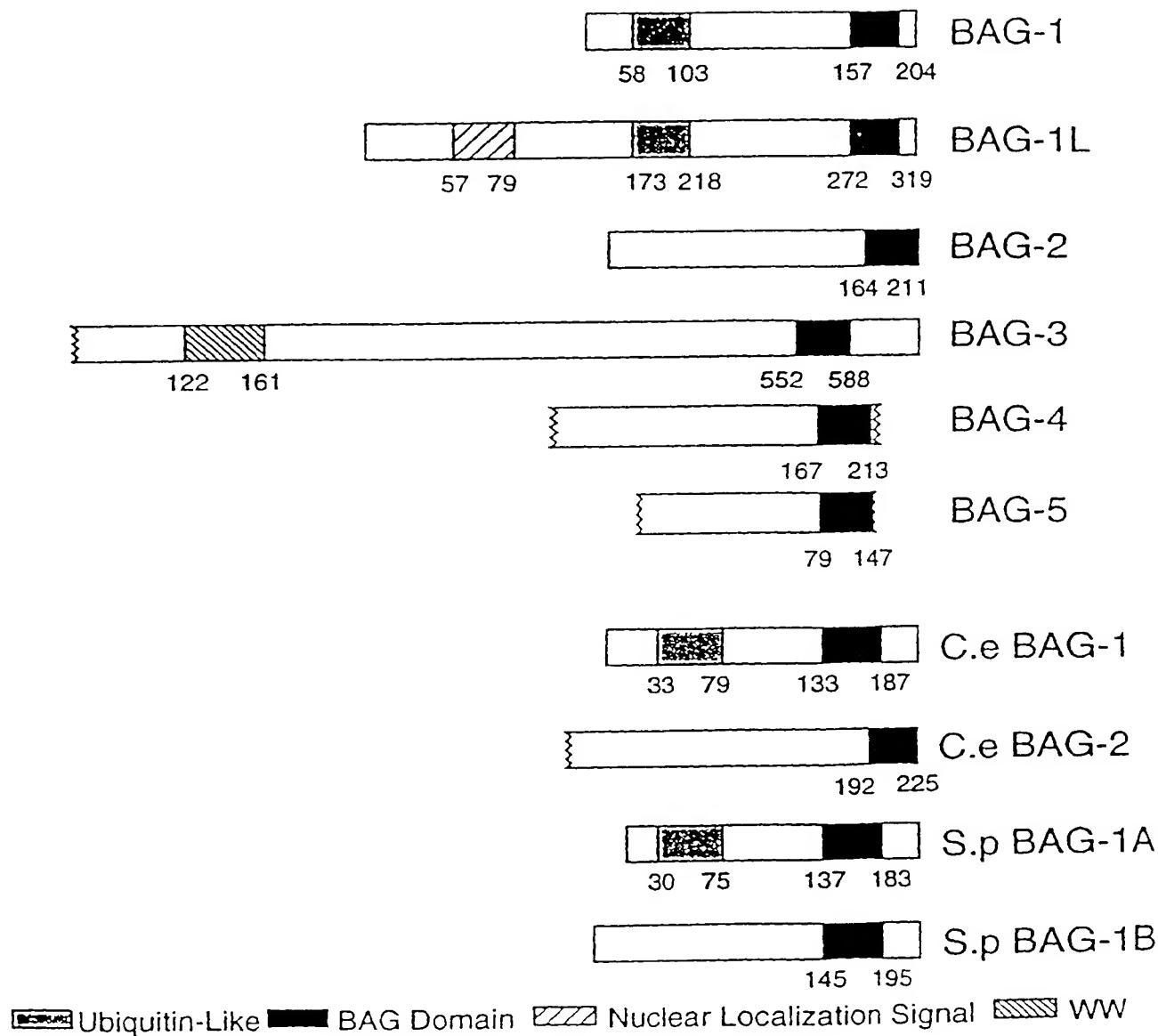


FIGURE 10B

157	C KLD KRE VKAT I QF N I LEE IT E	I T E - N F K E S F RLK R K G L V N K V Q A F I	I
158	KK T DKK Y LM Y EY IT K ELL ELD	I D P E G R A - - - - -	HBG-1
159	KK T DKA Y WL LEE IT K ELL ELD	I D P E G R A - - - - -	HBG-3
160	KK T DKA Y WL LEE IT K ELL ELD	I D P E G R A - - - - -	HBG-4
161	NRE DKN Y I RLE FLL T KOLLE IT N	I D P E G R A - - - - -	HBG-5
79	C KLD KRE VKAT I QF N I LEE IT E	I D P E G R A - - - - -	HBG-1
134	KK K KKK Y K Y EN V N EAE THL E IT N	I D P E G R A - - - - -	C.e BKG-1
133	KK K KKK Y K Y EN V N EAE THL E IT N	I D P E G R A - - - - -	S.p BKG-1A
137	KK K KKK LM E FLL O Q L K L D G V D V L G S E	I D P E G R A - - - - -	S.p BKG-1B
145	K DVO DL HT RL F I LL A R I K I D D V N E D D P	I D P E G R A - - - - -	HBG-2
164	LED D K K K R R L I T L R N T E N S I K A I K I L E H S K G A G S K I	I D P E G R A - - - - -	C.e BKG-2
192	ADD K K K R R L I N I N S Q U E N A R T K A D L - - - - -	I D P E G R A - - - - -	R.F.N

FIGURE 11

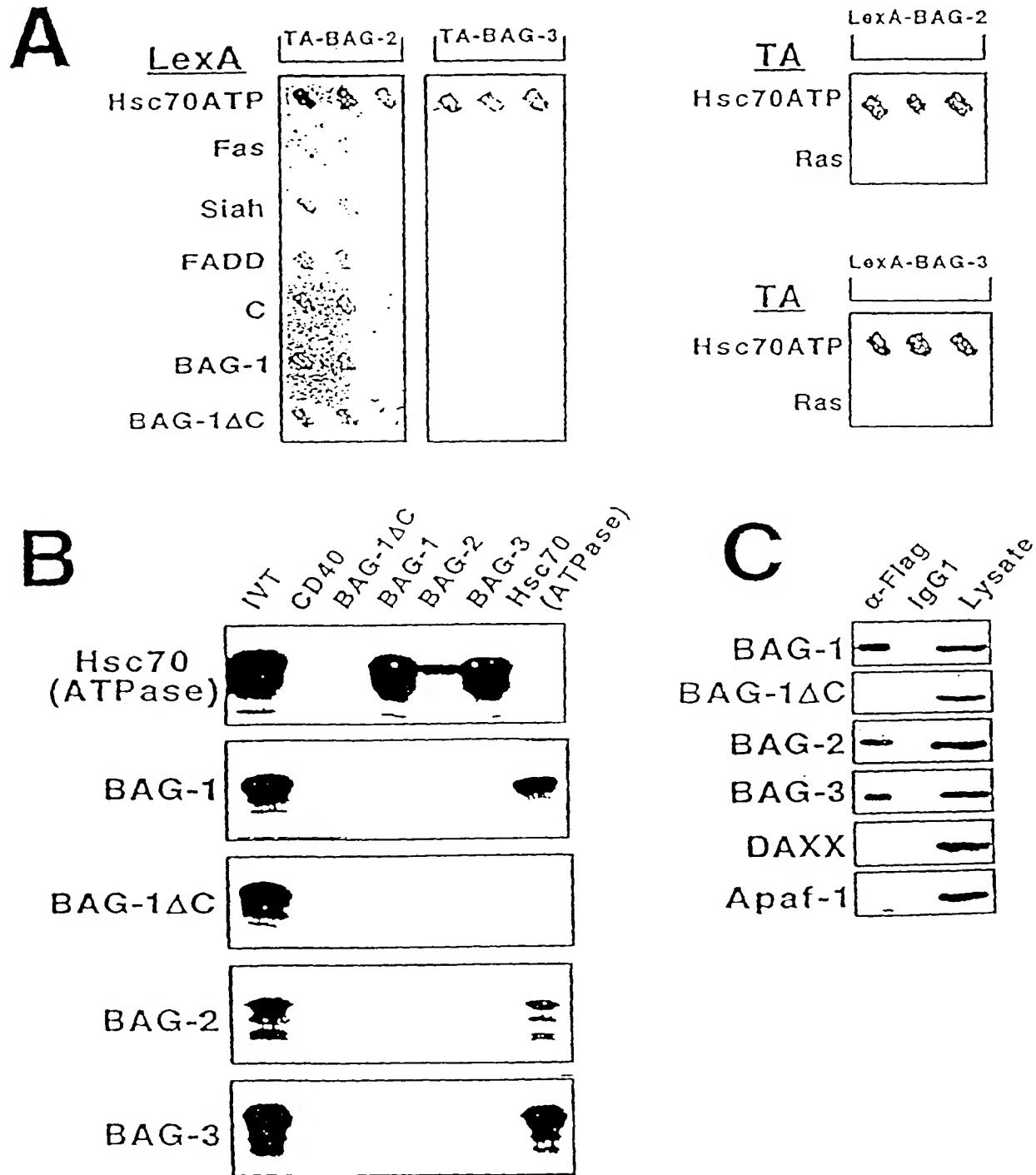


FIGURE 12

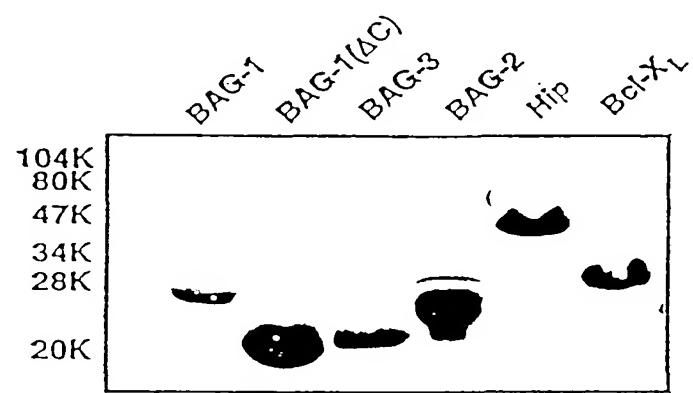


FIGURE 13

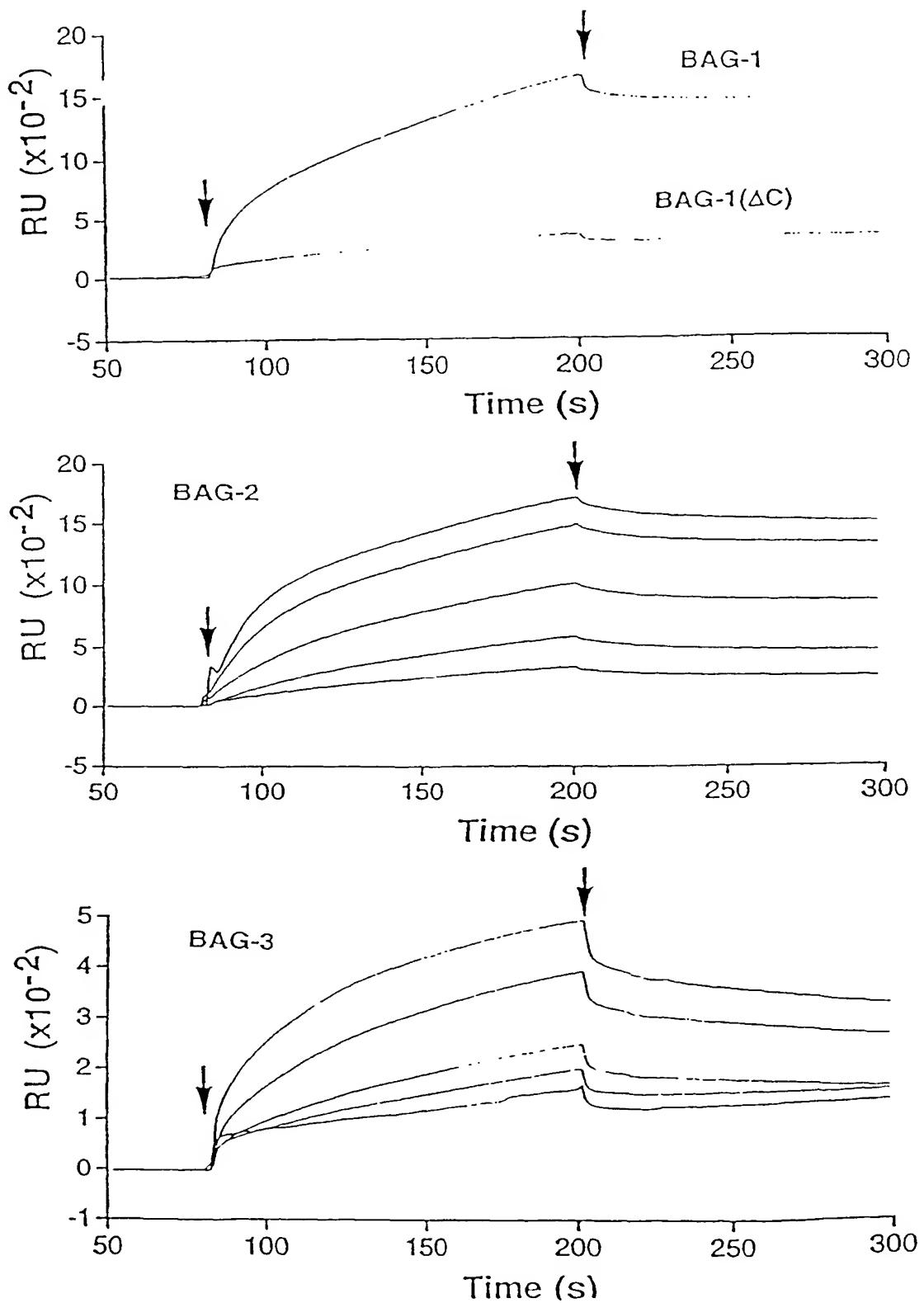


FIGURE 14

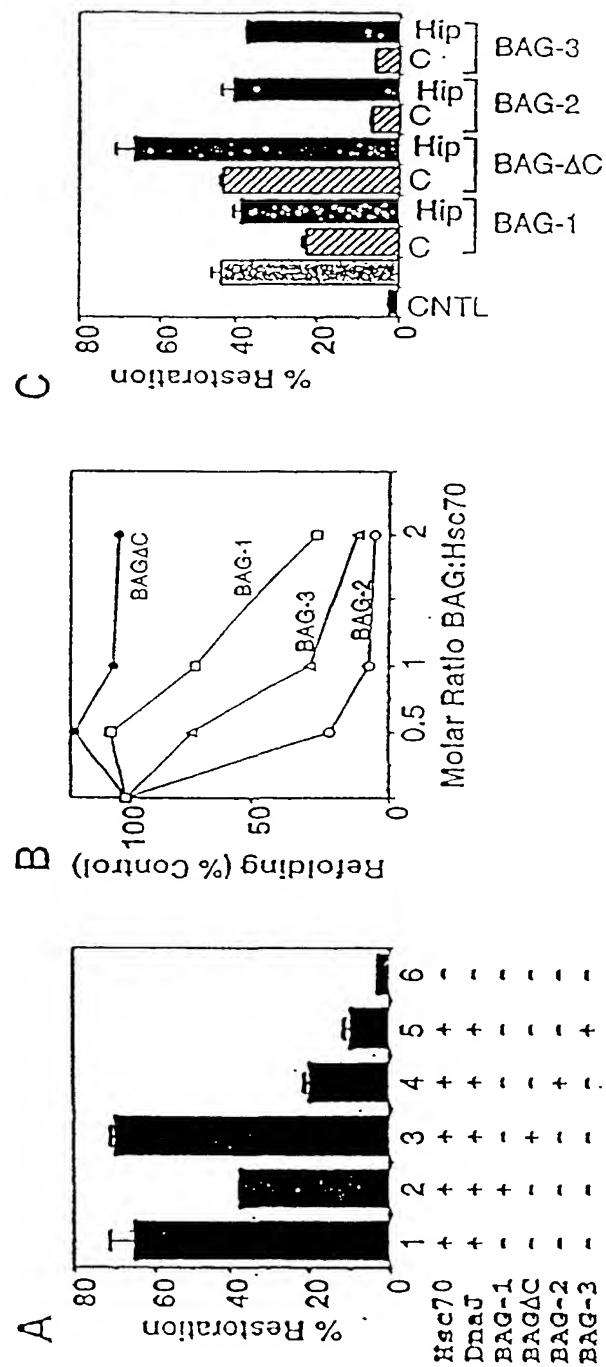


FIGURE 15A

50 GCGGAGCTCCGCATCCAACC CGGGGGGGG GCCAACTTCT CTGGACTTGGAA
CCAGAAGTTT CTAGCCGGCC AGTTGCTACC TCCCTTTTATC TCCCTCTTCC 100
CCTCTGGCAG CGAGGAGGCT ATTTCCAGAC ACTTCCACCC CTCCTCTGGCC 150
ACGTACCCCC CGCCTTTAAT TCATAAAGGT GCCCGGGGCC GGCTTCCGG 200
ACACGGTOGGC GGGGGAGGGG GGCCCCACGGC GGGGGGGG CCAGAGACTC 250
GGGGGGGA GCGAGGGGGC CGCAAGGGGG CGCGAGGGG CAGACCCCCA 300
CCCAGCATGA GCGGGGGCAC CGAACCTGGCCCG ATGATGGAGG TGGGGTCCGG 350
CAACGGTGCAC CGCGAACCCCTT TGCCCCCGG ATGGGAGATC AAGATCGACC 400
CGCAGAACCGG CTGGCCCTTC TTCTGTGGACC ACAACAGGGG CACCACTACG 450
TGGAAAGGCC CGCGGGTGCCTCTGTGGCC CGGGCAGGTG CACCCCTTCC ATGTCCTATCC 500
TGCCAAATGGC OCTTCCCGGG AGGGCTCTAG GCTGCCGGCT GCTAGGGAAAG 550
GCCACCCCTGT GTACCCCCAG CTCCGACCCAG GCTACATTC CATTCTGTG 600
CTCCATGAAG GCGCTGAGAA CGGGCAGGTG CACCCCTTCC ATGTCCTATCC 650
CCAGCCTGGG ATGCAGCGAT TCGGAACCTGA GGCGGCAGCA GCGGGCTCCTC 700
AGAGGGTCCCA GTCACCTCTG CGGGGCATATGC CAGAAACCAC TCAGGCCAGAT 750
AACAGTGTG GACAGGGTGGC AGGGGGGGC CCACCCAGC CCCCAAGCCTC 800
CCACGGGACCT GAGGGGGTCCC AGTCTCCAGC TGCCCTGTGAC TGCTCATCC 850
CATCCCTC GCCCAGGCTG CCTTCTCG GCAGGAGGAG CCTGGGGCAGT 900
CACCAAGCTCC CGCGGGGTA CATCTCCATT CGGGGTGATAC ACGAGCAGAA 950
CGTTACCCGG CCAGCAGGCC AGCCCTCTTT CCACAAAGGC CAGAAGACGC 1000
ACTACCCAGC GCAGAGGGGT GAGTACCGA CCCACCCAGCC TGTGTACAC 1050
AAGATCCAGG GGGATGACTG GGAGGCCCCGG CCCTCTGGGG CGGCATCC 1100
GTTCAAGGTCA TCTGTCCAGG GTGGATCGAG CGGGGGGGG TCACCAAGCCA 1150
GGAGCAGCAC GCCACTCCAC TCCOCCCTCGC CCATCCCGTGT GCACACCGTG 1200
GTCGACAGGC CTCAAGCAGGC CATGACCCAT CGAGAAACTG CACCTGTTC 1250
CCAGCCTGAA AACAAACCGA AAAGTAAGGC AGGGCCAGTT GGACCCAGAAC 1300
TCCCTCCTGG ACACATCCCA ATTCAAGTGA TCCGCAAAGA GGTGGATTCT 1350

FIGURE 15A

AACCTGTTT CCCAGAAGCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA 1400
 AGTCCCCCT GCTCCAGTTC CTTGCCCTC TOCCAGGCTT GGCCTCTG 1450
 CTGCCCCCT TTCCCCAAG AGTGTGGCTA CAGAAGAGAG GGCAGCCCCC 1500
 AGCACTGCC CTGCAGAAGC TACACCTCA AAACCAAGGAG AAGCCGGGGC 1550
 TCCCCAAA CATCCAGGAG TGCTGAAAGT GGAAGCCATC CTGGAGAAGG 1600
 TGAGGGGCT GGAGCAGGCT GTAGACAAC TTGAAGGCAA GAAGACTGAC 1650
 AAAAGTACCGT GAGTGTGAGAAGTATTTG ACCAAAGAGC TGCTGGCCCT 1700
 GGATTCAAGT GACCCCGAGG GACCGAGCGA TGTGCGTCAG GCCAGGAGAG 1750
 ACGGTGTCAG GAAGGTTCAAGT ACCATCTGG AAAAACATCTGA ACAGAAAGCC 1800
 ATTGATGTC CAGGTCAAGT CCAGGTCTAT GAACTCCAGC CCAGCAACCT 1850
 TGAGGCAAGT CAGCCACTGC AGGCAATCAT GGAGATGGGT GCCGTGGCAG 1900
 CAGACAAGGG CAAGAAAAAT GCTGAAATG CAGAAGATCC CCACACAGAA 1950
 ACCCAGCAGC CAGAAGCCAC AGCAGCAGCG ACTTCAAACC CCAGCAGCAT 2000
 GACAGACACC CCTGGTAACC CAGCAGCACC GTAGCCTCTG CCCTGTAATAA 2050
 ATCAGACTCG GAACCGATGT GTGCITTAAGG GAATTAAAG TTGCATGCAT 2100
 TTCAAGACT TTAAAGTCAGT TGGTTTTAT TAGCTGCTTG GTATGCAGTA 2150
 ACTTGGTGG AGGCAAACA CTAATAAAAG GGCTAAAAAG GAAAATGATG 2200
 CTTTCTTCT ATATTCCTAC TCTGTACAAA TAAAGAAGTT GCTTGTGTT 2250
 TGAGAAGTTT AACCCCCGTTG CTTGGTCTGC AGCCCTGTCT ACITGGGCAC 2300
 CCCCACCAACC TGTAGCTGT GGTGTGCACT GTGCTTTGT AGCTCTGGAC 2350
 TGGAGGGTA GATGGGGAGT CAATTACCA TCACATAAAAT ATGAAACATT 2400
 TATCAGAAAT GTGCCATT TAATGAGATG ATTTCTTCA TCTCATAAATT 2450
 AAAATAACCTG ACCTTAGAGA GAGTAAAATG TGCCAGGAGC CATAGGAATA 2500
 TCTGTATGTT GGATGACATT AATGCTACAT TTTC 2534

FIGURE 15B

MSAATHSPMM QVASGNGDRD PLPPGWEIKI DPQTGWPFFV DHNSRFTTTWN 50
DPRVPSEGPK ETPSSANGPS REGSRI PPAR EGHPVYVQLR PGYIPIVPLH 100
EGAENRQVHP FHVYQPQGMQ RFRTEAAAAA PQRSQSPLRG MPETTQPDQ 150
CGQVAAAAA QPPASHGPER SQSPAASDCS SSSSSASLPS SGRSSLSGHQ 200
LPRAGYISIPV IHEQNVTTPA AQPSFHKAQK THYPAQRGEY QTHQPVYHKI 250
QGDDWEPRPL RAASPFRSSV QGASSREGSP ARSSTPLHSP SPIRVHTVVD 300
RPQQPMTHRE TAPVSQOPENK PESKPGPGVGP ELPPGHIPIQ VIRKEVDSKP 350
VSQKPPPSE KVEVKVPPAP VPCPPPSGP SAVPSSSPKSVA TEERAAAPST 400
APAEAATPPKP GEAAPPKHP GVLKVEAILE KVQGLEQAVD NFEKGKTDKK 450
YLMIEEYLTK ELLALDSVDP EGRADVROAR RDGVRKVQTI LEKLEQKAID 500
VPGQVQVYEL QPSNLEADQP LQAIMEMGAV AADKGKKNAG NAEDPHTETQ 550
QPEATAAATS NPSSMTDTPG NPAAP
575

FIGURE 15C

CGGGGCTTC CGATCAGAAC CGGGGGGGG CGGACTTCTT CTGACTGAG CGAGAATTGTT CTGAGGGCC AGTTGCTTCC TOCCCTTTC
 TCTCTTCC CTCTTGACG CGAGGAGCT ATTGGAGAC ATCTGGCC CTCTCTGGC AGCTAGCC CGCTTTHAT TOCTTAACT
 CGGGGGGGG CGCTTGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 H S A T E S Z H M Q Y A S C H C S

 CGGGGGCTT TGGGGGGGGG ATGGACATC AGAGTGGAC CGGGGGGGG CTGGGGGGG TTGGGGGG CGGGGGGG CGGGGGGG
 Z D T L E S G E T T E T T E T T E T T E H S A T T

 TGGGGGGG CGGGGGGG CGCTTGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 V H D Z E V E S G E T T E T T E T T E H S A T C S L P

 CGGGGGGG CGGGGGGG CGCTTGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 A R E G E Z V T G L R E G T I T I T V L H E G A E H Z Q V

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 H T I E V T T Q Z C H Q R I E T Z A R A A A T Q E S Q S P L

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 R C M T E T T Q Z C H Q R I E T Z A R A A A T Q E S Q S P L

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 E A M T E T T Q Z C H Q R I E T Z A R A A A T Q E S Q S P L

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 E X J Q S Z A S S C S S S A S L E S S C E S S L C

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 R Q L Z A G I T S I T V E K I Q H V T R Z A R Q E S I E K A

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 Q E T H T Z A Q C I T Q T E Q I V T E K I Q C D D V T Z K

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 E L Z A R S T I S S V Q C R S S Z E C S T A R S S T P L R

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 S Z S I E V K T V V B Z I Q Q I H T E R I T A Z V S Q Z I

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 H E Z I S K E C V T G L L Z Z E K I Z I Q V I E K E V D

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 K I V S Q E Z Z E K E V K V E K V T Z Z A V Z C P Z E S Z

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 C I S A V Z S Z E S V R T Z Z A R A Z S T A Z A E R T Z

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 R I G Z A E R Z E K E I V L E V E A I L E V Y Q G L E Q R

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 V D H T E C K K T D E K T L H I Z E T L T E E L A Z S S V

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 D Z E C R A S V R Q A R E C V A E V Q T I L E K E E Q R A

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 I B V Z C Q T Q V T E L Q T E K L E A B Q Z L Q A E M E S

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 R V A R A D E C E K H A S H A Z D E E T I T Q E Q Z E A T A R A

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 T S H Z S S M T T T Z G H Z A A T

 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2140
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2250
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2340
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2430
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2520
 CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG CGGGGGGG
 2534

FIGURE 16A

50
CCTGAGGGC TCGGGCTACG GCCCCAGTGA CGTCCGTTACGGGGCT 100
ACTACGGGCCTGGGGCTGGGA GATGTGGGGTACACCCACCTCCACCTTA 150
TATCCTCTTGCCTGAACCTCCCCAGCCTCCCATTTCCCTGGGGGTGCG 200
CGGGGGGGCGGGAGAACACCTGGCTGGAGAAGGGGG 250
ATGGCTACTATCCCTGGGA GGCGGCTGGC CAGAGCCTGG TCGAGCCGG 300
GGAAGGCCAGGAGCAGCC ACCATATCCCT AGCTACAACTCTAATTG 350
GAATTCTACT GCGAGATCTA GGGCTCCCTAACCAAGTACA TATCTGTAA 400
GACCGAAATT GCAAGGGCCAG AGTTTGAATT CTTATACAAA TGGAGCGTAT 450
GGTCCAACAT ACCCCCCAGG CCCTGGGGCA AATACTGCCT CATACTCAGG 500
GGCTTATTAT GCACCTGGTT ATACTAGAC CAGTTACTCCACAGAAAGTTTC 550
CAAGTACTTA CGGTTCACTGGCAACAGCC CAACTCCAGTCTCTCGTTGG 600
ATCTATCCCCAGCAGGAACTG TCAGACTGAA GCACCCCCCTCTAGGGGGCA 650
GGTCCAGGA TATCCGGCCTT CACAGAACCC TGGAAATGACC CTGGCCCCATT 700
ATCCTTATGG AGATGGTAAT CGTAGTGTTC CACAATCAGG ACCGACTGTA 750
CGACCCACAAG AAGATGGCGTG GGCTTCTCCTGGTGCCTTATG GAATGGGTGG 800
CGGTTATCCC TGGCCCTCAT CAGGGCCCTAGCACCACCCGGGAATCTCT 850
ACATGACTGA AAGTACTTCA CCATGGCCTA GCAGTGGCTCTCCAGTCA 900
CCCCCTTCA CCCCCAGTCCA GCAGCCCAAG GATTCTTCATACCCCTATAG 950
CCAAATCAGAT CAAAGCATGA ACCGGCACAA CTTTCCCTTGCAGTGTCCATC 1000
AGTACGAATC CTGGGGGACA GTGATCAATG AAGATTCAAGA TCTTTGGAT 1050
TCCAAGTTCAGTATAGTGC TGAGCCTCAG CTGTATGGTA ATGCCACCAAG 1100
TGACCATCCC AACAAATCAAG ATCAAAGTAG CAGTCTTCCCTGAAGAATGTG 1150
TACCTTCAGA TGAAAAGTACT CCTCCGAGTA TTAAAAAAATCATACATGTG 1200
CTGGAGAAGG TCCAGTATCT TGAAACAAGAA GTAGAAGAAT TTGTAGGAAA 1250
AAAGACAGAC AAAGCATACT GGCTCTGGAGAAGAAATGCTA ACCAAGGAAC 1300

FIGURE 16A

TTTGGAACT GGATTCAAGTT GAAAAGTGGGG GCGAGGACTC TGTACGGCAG 1350
GCCAGAAAAG AGGCTGTTTG TAAGATTCAAG GCCATACTGG AAAAATTAGA 1400
AAAAAAAGGA TTATGAAAGG ATTAGAACAA AAGTGGAAAGC CTGTTACTAA 1450
CTTGACCAAA GAACACTTGA TTAGGTTAAAT TACCCCTCTTT TTGAAATGCC 1500
TGTGATGAC AAGAAGCAAT ACATTCAGC TTTCCCTTG ATTTTATACT 1550
TGAAAAGTGC GCAAAGGAAT GGAAGAAAT TTTAGTCATG AAGTTGTTT 1600
CAGTTTCAGA CGAATGAATG TAATAGAAA CTATGGAGTT ACCAAATATTG 1650
CCAAGTAGAC TCACTCCTTA AAAAATTAT GGATATCTAC AAGCTGCTTA 1700
TTACAGCAG GAGGGAAACA CACTTCACAC AACAGGGCTTA TCAGAAAACCT 1750
ACCAGATGAA ACTGGATATA ATTTGAGACA AACAGGGATGT GTTTTTAA 1800
ACATCTGGAT ATCTTGTCACTTTTGAC ATTGTGACTG CTTCAACAT 1850
ATACCTTCATG TGTAATTATA GCTTAGACCT TAGCCTCTT GGACTTCTGT 1900
TTGGTTTGT TATTTGCAGT TTACAAATAT AGTATTATC TCTAAAAA 1950
AAAAAAAGAA AAAAAGA 1966

FIGURE 16B

MSALRRSGYGPSPDGPSYGRYYGPGGGDVPVHPPPLYLRPEPPOPPISWRVRGGGPAETTLGECCCCDGYYPPSGGAWP
EPGRAGGSHDEQPPPPSYNSNMWNSTARSRAPYPSYSTVRPELQGOSLNSYNTNGAYGPTYPGPAGNTASYSGAYYAPGY
TQTSYSTEMPSYRSGNSPPTPVSRLWYPQDCQTEAPPLRGQVPGYPPSQNPGMTLPHYHYGDGNRNSVPQSGPTVRPQE
DAWASPGAYGMGGRYPWPSASSAPSAPPGNLYMTESTSPWPSSGSPQSPSPPVQQPKDSSYPPYSQSDQSMNRHNIFPCSVHQ
YESSGTVINEDSDLLDSQVQYSAEPQLYGNATSDHPNNQDQSSSLPEECVPSDESTPPSIKKIHLERKVQYLEQEWEFF
VGKKTDKAYWLLEEMLTKELEELDSVETGGQDSVRQARKEAVCKIQAILEKLEKKGL

FIGURE 16C

FIGURE 17A

CCCCCCCCCC CCGAAGAAG CCGGGAGCGG CTGGCTGCAGC 50
 CAGTAGGGGC CCCTTACCG GCTGCCCGC TCAGACCTAG TGGGAGGG 100
 TGCCAGGGCAT GCAGCTGGG GCCCCAGCTCC GGTCGGCAC CCCGTAAGG 150
 GCTGATCTTC CACCTGCCA CCTCAGCCAC GGACGCCAA GACCCGATCC 200
 AATTCAAGCT TCTTTGGTG CTTGTGAAAC TGAACACAAC AAAAGTATGG 250
 ATATGGAAA CCAACATCCT TCTATTAGTA GGCTTCAGGA AAATCCAAAAG 300
 GAAGTAAAAA GTGTAGAACCA GCAAGTTATC GGCTTCAGTG GTCTGTCAGA 350
 TGACAAGAAT TACAAGAAAAC TGGAGAGGAT TCTAACAAAAA CAGCTTTTG 400
 AAATAGACTC TGTAGATACT GAAGGAAAAG GAGATATTCA GCAAGGCTAGG 450
 AAGGGGGCAG CACAGGGAGAC AGAACGTCCT CTCAAAGAGT TGGAGGCAGAA 500
 TGCAAACAC CCACACCGGA TTGAAATACA GAACATTTT GAGGAAGCCC 550
 AGTCCCTCGT GAGAGAGAAA ATTGTGCCAT TTTATAATGG AGGCAACTGC 600
 GTAAGTGTG AGTTTGAGA AGGCATCCAA GATATCATTG TGAGGGCTGAC 650
 ACATGTTAAA ACTGGAGGAA AAATCTCCTT GCGGAAAGCA AGGTATCACA 700
 CTTTAACCAA AATCTGTGCG GTGCAAGAGA TAATCGAAGA CTGCATGAAA 750
 AAGCAGCCCT CCCTGCCGCT TTCCCGAGGAT GCACATCCTT CCGTTGCCAA 800
 AATCAACTC GTGATGTGTTG AGGTGAACAA GGCCCCGAGGG GTCCCTGATTG 850
 CACTTCTGAT GGGTGTGAAC ACAATGAGA CCTGCAGGCA CTTATCCTGT 900
 GTGCTCTCGG GGCTGATCGC TGACCTGAT GCTCTAGATG TGTGGGGCCG 950
 GACAGAAATC AGAAATTATC GGAGGGGGT AGTAAAGAT ATCAACAAAT 1000
 TATTGAAATA TCTGGATTTG GAAGAGGAAG CAGACACAAAC TAAAGCATT 1050
 GACCTGAGAC AGAATCATTG CATTAAAGG ATAGAAAAGG TCCTCAAGAG 1100
 AATGAGAGAA ATAAAAATG AACTTCCTCCA AGCACAAAAC CCTTCTGAAAT 1150
 TGTACCTGAG CTCCAAACAA GAATTGCGAGG GTTTAATTGG ACAGTTGGAT 1200
 GAGGTAAGTC TTGAAAAAAA CCCCTGCATC CGGGAAGCCA GGAGAAGAGC 1250
 AGTGATCGAG GTGCAAAACTC TGATCACATA TATTGACTTG AAGGGAGGCC 1300

FIGURE 17A

TTGAGAAAAG AAAGCTGTTT GCTTGTGAGG AGCACCACAT CCATAAAGCC 1350
 GTCTGGAACG TCCCTGGAAA CCTTGTGAG ATCCAGGGAG AAGTTCTTTC 1400
 ATTTGATGGA AATCGAACCG ATAAGAACTA CATCCGGCTG GAAGAGCTGC 1450
 TCACCAAGCA GCTGCTAGCC CTGGATGCTG TTGATCCGCA GGGGAAAGAG 1500
 AAGTGTAAAGG CTGCCAGGAA ACAAGCTGTG AGGCTTGCAG AGAATATTCT 1550
 CAGCTATCTC GACCTGAAAT CTGATGAAATG GGAGTACTGA AATACCAGAG 1600
 ATCTCACCTT TGATACCTGTT TTGCACTTCA TATGTTAGA 1650
 GAGCCTTCAG TTCAATTGATT TATACGTGCA TATTTCACTGTC TCAGTATTA 1700
 TGATTGAAGC AAATTCTATT CAGTATCTGC TGCTTTGAT GTTGCAAGAC 1750
 AAATATCATT ACAGGCACGTT AACTTTCCA TTGGGATCAT TATCTGTATG 1800
 ATGTGGGTG GTTGTGTTGG TTGGTCCCTT TTTTGGCTT TTTAATCAGA 1850
 AACACAAATA GAGGCAGCTT TTGTAGATT TAAATGGGTT GTGCAAGCAT 1900
 TAAAATGCG AGTCTTCAGA ATCTAGAACT AGGCATAACC TTACATAATA 1950
 CTAGGAAAAT TATGAGAAAG GGGAAATT TTGGTTAAATA AGAGTAAGGT 2000
 TCAAACACAA GCAGTACATG TTCTGTTCA TTATGCTCGA TAGAAGGGCTT 2050
 TTTTTTCACT TATAAGGCCT GATTGGTCCCT ACCCAGCCTT ACGGGGTGGG 2100
 GTTTTTTGT TTGTTCAAGAC AGTCTGTTCT TTGGTAAACA TTTTAGTTG 2150
 GAAAAAACAGGC ATCTGCATT TCCCCATCCT CTACGTTTA GAGAGGAATC 2200
 TTGTTTTGT GTGCAACATA AGAAAATTAT GAAAACCTAAT AGCCAAAAAA 2250
 CCTTTGAGAT TGCATTAAG AGAAGGGATA AAGGACCCAGC ATAATACCT 2300
 TGTAAGTTCG TTTGGTTGT AAAATCTGAG CTTATAGTTT TCCTTAGTGA 2350
 GTAAAATTCACT AAGGATGGGA ACATTTAAAT TAAGTTAATG GGCCCTTAA 2400
 AAAAAAAAG GAAACACTCA TACCTGTAGT TGGAGGATGA ATACTGGAGA 2450
 CGGGTTACCA ATGTCAGGTT ATACCTAAAC TAAATCAGAA AGTCTGAATG 2500
 TAGGCACATAA TGGTTCTCTT CTGTTGTCAA AGGCTGTAAA ATGGACAGCC 2550
 TTGTCACACC TCCCCGGTGC TGTTCACCAA CGTGAGGTA GACGGCTGTC 2600

FIGURE 17A

GTAAACCCAGA GGGACCAGGC CTTCCCTAGGGT TTTCTAGGCC GTCAGCTGTT 2650
 AACCACTCAC TTAGTAAATG TCATAACTAC ACTGTCTCCA GGACCAATCA 2700
 GTGAAACCTG CTCGGAATTA AAGGCTCCCT CTGGGTGCCT GCTGAACAAAC 2750
 TGAGCTCATG TCATGGCAT GTGGTGGTT CTCTGTTGCC TGAAAGAGGCC 2800
 ATAAAGTCA GTCGTTGCGTG AAGCATCTCT CTTCTAAAGG ATGTTGTTTT 2850
 CCATAAATGC TTCTCTGAGGA TCCGGTACAA AATGATTTC CAAAGTTCTG 2900
 AAGTGCCTTG AGAACATGTG GTTCCGAGTG TTATAACAGA CTCCCTCCCC 2950
 GGGTCACCTT TTGCTCTGGTC ATCCTGTTAG AGTACATCTT TGGAAATCCA 3000
 GGGTAATATT CTCTTTCAGA GATGCTCATT GTGTAACTCT GTGTAGGGAG 3050
 ATAGTCACTT TAAACAGCTC AAAGTAGCTA GCTAAAGGGAG TAGCCCTTAA 3100
 TACTCTAAAG ATGACAGAAAG CATAGCCCTT ACAAAATCTT CAGCTTGTCT 3150
 CTCAGTATT CCCAATCATG AAAATCCCTT GCTATGTCCT TCCTACTAGA 3200
 AATGTTCTAG AATCGCTGGA CGTGGGGTC AGAGGGCAGT CGGTATTTAG 3250
 GCGGTGAGCT TCCCATACTA CTGCGAGTCC AACTCCTGCC AACCGGGGC 3300
 TCAAGGCAGG TCATTGGAAT CCACGTTTTG GCCACAGTAG TTGTAGGGATT 3350
 GCTTTCTGT ATCATTAATT TAGAATGCTC TTAAAATCTT GAGGAAGAGT 3400
 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3450
 TGAGATGGAG TCTCTGTTGC CCAGGCTGCCA 3500
 GTGCAGTGGT GCCATCTCAG CTCAC TGCAA CCTCCACCTC CCAGGTTCAA 3550
 GCGATTCTCC TGCCCTCAGCC ACCTGAGTAG CTGGGAGTAC AGGCATGTGG 3600
 CACCATGCCCT GGCTAATTTT TGTATTTTTA ATAGAGTTGA GATTTCACCA 3650
 TGATGGTCAG GCTGGTCTCG AACTCCTGAC CTCGTGATCC GCCCCCCTCG 3700
 GCCCCCCAAA GTGCTGGGAT TAACGGGTGT GAGCCACGGC GCCCAGCCCA 3750
 GGAAAGAGTTT TTAAATTAGA GCTCTGTTTA ATTATACCAAC TGGGAAATCA 3800
 TGGTTACGCT TCAGGCATAT TCTTCCCCAG AGTACTACTT ACATTTAAA 3850
 TTTCATTTTG TAAAGTTAA TGTCAAGCATT CCCTTAAAA GTGTCCATTG 3900
 TTCTTTGAAA GTAGACGTTT CAGTCATTCT TTTCAAAACAA GTGTTTGTGT

FIGURE 17A

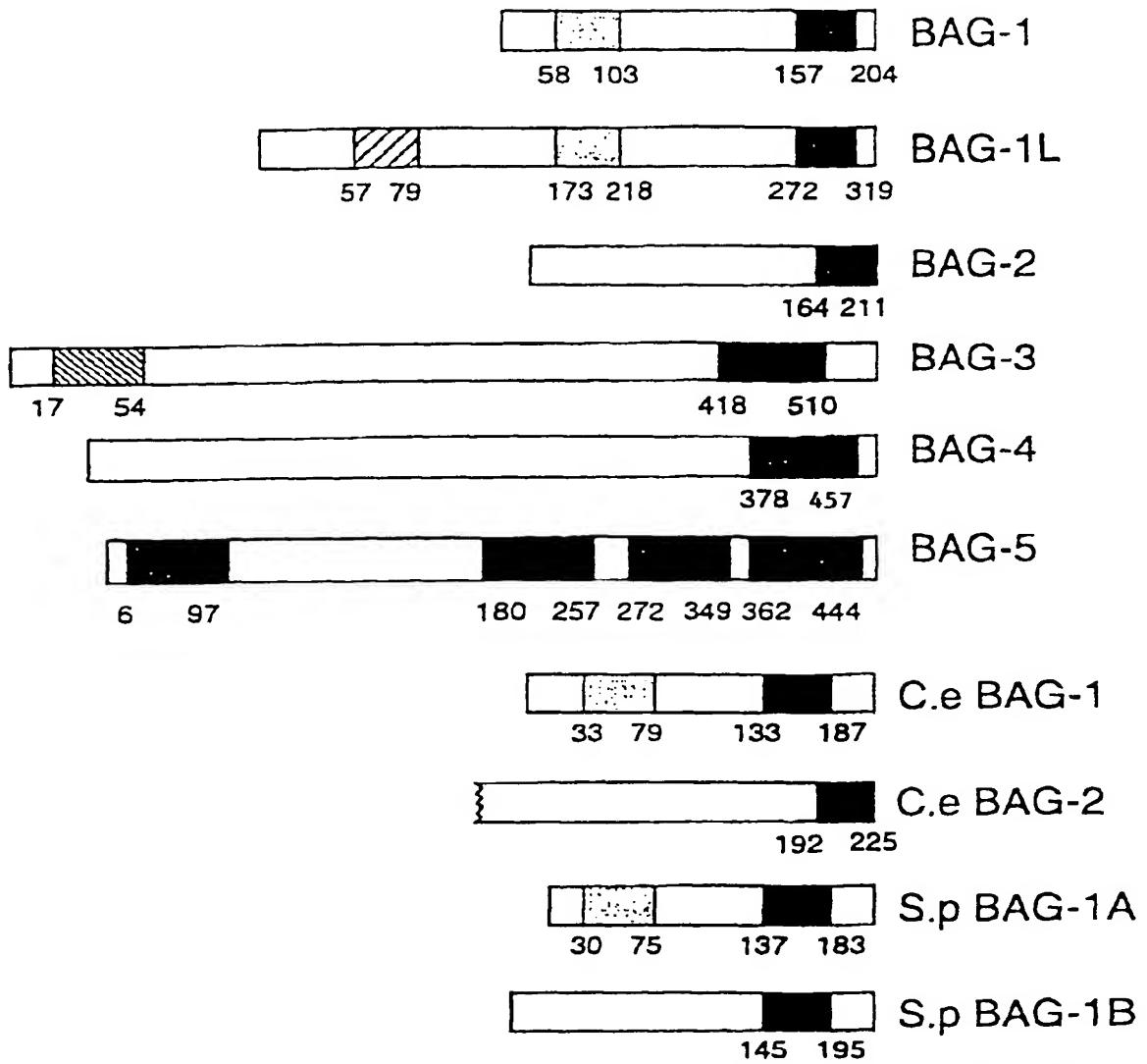
ACCTTTGCC AAGCTGTGGG CATCGTGTGT GAGTACAGGG TGCTCAGCTC
TTCCACCGTC ATTTTGAATT GTTCACATGG GTAATTGGTC ATGGAAATGA 4000
TCAGATTGAC CTTGATTGAC TGTCAAGGCAT GGCTTTGGTTT CTAGTTTCAA 4050
TCTGTTCTG TTCCCTGTAC CGGATTATTCT TACTCCTGCA ATGAAACCTG 4100
TTGACACCGG ATTTAGCTCT TGTGGGCCTT CGTGGGGAGC TGTGTGTGTT 4150
AATATGAGCT ACTGCATGTA ATTCTTAAAC TGGGCTTGTAC ATTTGTATT 4200
GTATTTTGT GATCTGTAAT GAAAAGAAC TGTACTGCA GTAAAACCTA 4250
CTCCCCAAA ATGTGTGGCT TTGGGTCTGC ATTAAACGCT GTAGTCCATG 4300
TTCATGCC 4308

FIGURE 17B

MDMGNQHPSI SRLQEIQKEV KSVEQQVIGF SGLSDDKNYK KLERLTIKQL
50
FEIDSVDTEG KGDIQQAQARKR AAQETERLILK ELEQNANHPH RIEIQNIFEE
100
AQSLVREKIV PFYNGNCVT DEFEEGIQDI ILRLTHVKTG GKISLRKARY
150
HTLTKICAVQ EIIEDCMKKQ PSIPLSEDAH PSVAKINFVM CEVNKARGVL
200
IALLMGVNNN ETGCRHLSCVL SGLIADLAL DVCGRTEIRN YRREVEDIN
250
KLLKYLDLEE EADTTKAFDL RQNHSILKIE KVLKFMREIK NELLQAQNPS
300
ELYLSSKTEL QGLIGOLDEV SLEKNPCIRE ARRRAVIEVQ TLITYIDLKE
350
ALEKRKLFAC EEEHPHSHKAWW NVLGNLSEIQ GEVLSFDGNR TDKNYIRLEE
400
LLTKQLLAD AVDPQGEEKC KAARKQAVRL AQNILSYLDL KSDEWEY
447

FIGURE 17C

FIGURE 18



■ Ubiquitin-Like ■ BAG Domain ■/■ Nuclear Localization Signal ■■■ WW

SEQUENCE LISTING

<110> Reed, John C.

Takayama, Shinichi

The Burnham Institute

<120> Novel BAG Proteins and Nucleic Acid Molecules Encoding
Them

<130> FP-LJ 3646

<140>

<141>

<150> 09/150,489

<151> 1998-09-09

<160> 24

<170> PatentIn Ver. 2.0

<210> 1

<211> 1291

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (46)..(1080)

<400> 1

acgcccgcgt cagcttccat cgctggggcg tcaacaagtgcgggcctg gct cag cgc 57
Leu Ala Gln Arg
1

ggg ggg gcg aga ccg cga ggc gac cgg gag cgg ctg ggt tcc cgg 105
Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg Leu Gly Ser Arg
5 10 15 20

ctg cgc gcc ctt cgg cca ggc cgg gag cgc cag tcg gag ccc cgg 153
Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln Ser Glu Pro Pro
25 30 35

gcc cag cgt ggt ccg cct ccc tct cgg cgt cca cct gcc cgg agt act 201
Ala Gln Arg Gly Pro Pro Ser Arg Arg Pro Pro Ala Arg Ser Thr
40 45 50

gcc agc ggg cat gac cga ccc acc agg ggc gcc gcc ggc gct cgc 249

Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala Ala Gly Ala Arg			
55	60	65	
agg ccg cgg atg aag aag aaa acc cgg cgc cgc tcg acc cgg agc gag 297			
Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser Thr Arg Ser Glu			
70	75	80	
gag ttg acc cgg agc gag gag ttg acc ctg agt gag gaa gcg acc tgg 345			
Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu Glu Ala Thr Trp			
85	90	95	100
agt gaa gag gcg acc cag agt gag gag gcg acc cag ggc gaa gag atg 393			
Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln Gly Glu Glu Met			
105	110	115	
aat cgg agc cag gag gtg acc cgg gac gag gag tcg acc cgg agc gag 441			
Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser Thr Arg Ser Glu			
120	125	130	
gag gtg acc agg gag gaa atg gcg gca gct ggg ctc acc gtg act gtc 489			
Glu Val Thr Arg Glu Glu Met Ala Ala Gly Leu Thr Val Thr Val			
135	140	145	
acc cac agc aat gag aag cac gac ctt cat gtt acc tcc cag cag ggc 537			
Thr His Ser Asn Glu Lys His Asp Leu His Val Thr Ser Gln Gln Gly			
150	155	160	
agc agt gaa cca gtt gtc caa gac ctg gcc cag gtt gaa gag gtc 585			
Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val Val Glu Glu Val			
165	170	175	180
ata ggg gtt cca cag tct ttt cag aaa ctc ata ttt aag gga aaa tct 633			
Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe Lys Gly Lys Ser			
185	190	195	
ctg aag gaa atg gaa aca ccg ttg tca gca ctt gga ata caa gat ggt 681			
Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly Ile Gln Asp Gly			
200	205	210	
tgc cgg gtc atg tta att ggg aaa aag aac agt cca cag gaa gag gtt 729			
Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro Gln Glu Glu Val			
215	220	225	
gaa cta aag aag ttg aaa cat ttg gag aag tct gtg gag aag ata gct 777			
Glu Leu Lys Lys Leu Lys His Leu Glu Lys Ser Val Glu Lys Ile Ala			
230	235	240	
gac cag ctg gaa gag ttg aat aaa gag ctt act gga atc cag cag ggt 825			

Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly Ile Gln Gln Gly			
245	250	255	260
ttt ctg ccc aag gat ttg caa gct gaa gct ctc tgc aaa ctt gat agg 873			
Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys Lys Leu Asp Arg			
265	270	275	
aga gta aaa gcc aca ata gag cag ttt atg aag atc ttg gag gag att 921			
Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile Leu Glu Glu Ile			
280	285	290	
gac aca ctg atc ctg cca gaa aat ttc aaa gac agt aga ttg aaa agg 969			
Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser Arg Leu Lys Arg			
295	300	305	
aaa ggc ttg gta aaa aag gtt cag gca ttc cta gcc gag tgt gac aca 1017			
Lys Gly Leu Val Lys Val Gln Ala Phe Leu Ala Glu Cys Asp Thr			
310	315	320	
gtg gag cag aac atc tgc cag gag act gag cgg ctg cag tct aca aac 1065			
Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu Gln Ser Thr Asn			
325	330	335	340
ttt gcc ctg gcc gag tgaggtgttag cagaaaaagg ctgtgctgcc ctgaagaatg 1120			
Phe Ala Leu Ala Glu			
345			
gccccaccag ctctgccgtc tctggatcgg aatttacctg atttctttag ggctgctggg 1180			
ggcaactggc catttgccaa ttttctact ctcacactgg ttctcaatga aaaatagtgt 1240			
ctttgtgatt tgagtaaagc tcctattctg ttttcacaa aaaaaaaaaa a 1291			

<210> 2
 <211> 345
 <212> PRT
 <213> Homo sapiens

<400> 2			
Leu Ala Gln Arg Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg			
1	5	10	15
Leu Gly Ser Arg Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln			
20	25	30	
Ser Glu Pro Pro Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro			
35	40	45	

Ala Arg Ser Thr Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala
50 55 60

Ala Gly Ala Arg Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser
65 70 75 80

Thr Arg Ser Glu Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu
85 90 95

Glu Ala Thr Trp Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln
100 105 110

Gly Glu Glu Met Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser
115 120 125

Thr Arg Ser Glu Glu Val Thr Arg Glu Glu Met Ala Ala Ala Gly Leu
130 135 140

Thr Val Thr Val Thr His Ser Asn Glu Lys His Asp Leu His Val Thr
145 150 155 160

Ser Gln Gln Gly Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val
165 170 175

Val Glu Glu Val Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe
180 185 190

Lys Gly Lys Ser Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly
195 200 205

Ile Gln Asp Gly Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro
210 215 220

Gln Glu Glu Val Glu Leu Lys Lys Leu Lys His Leu Glu Lys Ser Val
225 230 235 240

Glu Lys Ile Ala Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly
245 250 255

Ile Gln Gln Gly Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys
260 265 270

Lys Leu Asp Arg Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile
275 280 285

Leu Glu Glu Ile Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser
290 295 300

Arg Leu Lys Arg Lys Gly Leu Val Lys Lys Val Gln Ala Phe Leu Ala
305 310 315 320

Glu Cys Asp Thr Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu
325 330 335

Gln Ser Thr Asn Phe Ala Leu Ala Glu
340 345

<210> 3

<211> 1179

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (160)..(792)

<400> 3

gcagccgcgg tgtcgcaag tcctccggg ttgccccgc ggcgtcagag ggagggcggg 60

cgccgcgttg gtgacggcga ccctgcagcc caaggagcgc tccactcgct gccgcccggag 120

ggccgggtgac ctcttggcta ccccgcgctcg gaggcttag atg gct cag gcg aag 174
Met Ala Gln Ala Lys
1 5

atc aac gct aaa gcc aac gag ggg cgc ttc tgc cgc tcc tcc atg 222
Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys Arg Ser Ser Ser Met
10 15 20

gct gac cgc tcc agc cgc ctg ctg gag agc ctg gac cag ctg gag ctc 270
Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu Asp Gln Leu Glu Leu
25 30 35

agg gtt gaa gct ttg aga gaa gca gca act gct gtt gag caa gag aaa 318
Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala Val Glu Gln Glu Lys
40 45 50

gaa atc ctt ctg gaa atg atc cac agt atc caa aat agc cag gac atg 366
Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln Asn Ser Gln Asp Met
55 60 65

agg cag atc agt gac gga gaa aga gaa gaa tta aat ctg act gca aac 414
Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu Asn Leu Thr Ala Asn
70 75 80 85

cgt ttg atg gga aga act ctc acc gtt gaa gtg tca gta gaa aca att 462
 Arg Leu Met Gly Arg Thr Leu Thr Val Glu Val Ser Val Glu Thr Ile
 90 95 100

aga aac ccc cag cag caa gaa tcc cta aag cat gcc aca agg att att 510
 Arg Asn Pro Gln Gln Glu Ser Leu Lys His Ala Thr Arg Ile Ile
 105 110 115

gat gag gtg gtc aat aag ttt ctg gat gat ttg gga aat gcc aag agt 558
 Asp Glu Val Val Asn Lys Phe Leu Asp Asp Leu Gly Asn Ala Lys Ser
 120 125 130

cat tta atg tcg ctc tac agt gca tgt tca tct gag gtg cca cat ggg 606
 His Leu Met Ser Leu Tyr Ser Ala Cys Ser Ser Glu Val Pro His Gly
 135 140 145

cca gtt gat caq aag ttt caa tcc ata gta att ggc tgt gct ctt gaa 654
 Pro Val Asp Gln Lys Phe Gln Ser Ile Val Ile Gly Cys Ala Leu Glu
 150 155 160 165

gat cag aag aaa att aag aga aga tta gag act ctg ctt aga aat att 702
 Asp Gln Lys Lys Ile Lys Arg Arg Leu Glu Thr Leu Leu Arg Asn Ile
 170 175 180

gaa aac tct gac aag gcc atc aag cta tta gag cat tct aaa gga gct 750
 Glu Asn Ser Asp Lys Ala Ile Lys Leu Leu Glu His Ser Lys Gly Ala
 185 190 195

ggt tcc aaa act ctg caa caa aat gct gaa agc aga ttc aat 791
 Gly Ser Lys Thr Leu Gln Gln Asn Ala Glu Ser Arg Phe Asn
 200 205 210

tagtcttcaa acctaagagc atttacacaa tacacaagggt gtaaaaatga taaaatacta 852

ttttaattga taactagttc tttgttaggt ataaccactt agttgacact gatagttgtt 912

tcagatgagg aaaatattcc atcaagtatc ttcatgtttt tgaataacaa aactagcaat 972

attttaatta tctatctaga gattttttag attgaattct tgtcttgtac taggatctag 1032

catatttcac tattctgtgg atgaatacat agtttgtggg gaaaacaaac gttcagctag 1092

gggcaaaaag catgactgct tttccctgtc tggcatggaa tcacgcagtc accttggca 1152

ttttagtttac tagaaattct ttactgg 1179

<210> 4
<211> 211
<212> PRT
<213> Homo sapiens

<400> 4

Met	Ala	Gln	Ala	Lys	Ile	Asn	Ala	Lys	Ala	Asn	Glu	Gly	Arg	Phe	Cys
1					5				10					15	
Arg	Ser	Ser	Ser	Met	Ala	Asp	Arg	Ser	Ser	Arg	Leu	Leu	Glu	Ser	Leu
	20							25					30		
Asp	Gln	Leu	Glu	Leu	Arg	Val	Glu	Ala	Leu	Arg	Glu	Ala	Ala	Thr	Ala
	35					40				45					
Val	Glu	Gln	Glu	Lys	Glu	Ile	Leu	Leu	Glu	Met	Ile	His	Ser	Ile	Gln
	50					55				60					
Asn	Ser	Gln	Asp	Met	Arg	Gln	Ile	Ser	Asp	Gly	Glu	Arg	Glu	Glu	Leu
	65				70				75			80			
Asn	Leu	Thr	Ala	Asn	Arg	Leu	Met	Gly	Arg	Thr	Leu	Thr	Val	Glu	Val
	85					90				95					
Ser	Val	Glu	Thr	Ile	Arg	Asn	Pro	Gln	Gln	Gln	Glu	Ser	Leu	Lys	His
	100					105				110					
Ala	Thr	Arg	Ile	Ile	Asp	Glu	Val	Val	Asn	Lys	Phe	Leu	Asp	Asp	Leu
	115					120				125					
Gly	Asn	Ala	Lys	Ser	His	Leu	Met	Ser	Leu	Tyr	Ser	Ala	Cys	Ser	Ser
	130					135				140					
Glu	Val	Pro	His	Gly	Pro	Val	Asp	Gln	Lys	Phe	Gln	Ser	Ile	Val	Ile
	145					150				155			160		
Gly	Cys	Ala	Leu	Glu	Asp	Gln	Lys	Lys	Ile	Lys	Arg	Arg	Leu	Glu	Thr
	165					170				175					
Leu	Leu	Arg	Asn	Ile	Glu	Asn	Ser	Asp	Lys	Ala	Ile	Lys	Leu	Leu	Glu
	180							185			190				
His	Ser	Lys	Gly	Ala	Gly	Ser	Lys	Thr	Leu	Gln	Gln	Asn	Ala	Glu	Ser
	195					200				205					
Arg	Phe	Asn													
	210														

<210> 5
 <211> 2528
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)...(2031)

<400> 5

gct	gag	ctc	cgc	atc	caa	ccc	cgg	gcc	gct	gcc	aac	ttc	tct	gga	ctg	48
Ala	Glu	Leu	Arg	Ile	Gln	Pro	Arg	Ala	Ala	Ala	Asn	Phe	Ser	Gly	Leu	
1		5			10						15					

gac	cag	aag	ttt	cta	gcc	ggc	cag	ttg	cta	cct	ccc	ttt	atc	tcc	tcc	96
Asp	Gln	Lys	Phe	Leu	Ala	Gly	Gln	Leu	Leu	Pro	Pro	Phe	Ile	Ser	Ser	
			20					25					30			

ttc	ccc	tct	ggc	agc	gag	gag	gct	att	tcc	aga	cac	ttc	cac	ccc	tct	144
Phe	Pro	Ser	Gly	Ser	Glu	Glu	Ala	Ile	Ser	Arg	His	Phe	His	Pro	Ser	
			35					40			45					

ctg	gcc	acg	tca	ccc	ccg	cct	tta	att	cat	aaa	ggt	gcc	cg	ccg	192	
Leu	Ala	Thr	Ser	Pro	Pro	Pro	Leu	Ile	His	Lys	Gly	Ala	Arg	Arg	Arg	
			50				55			60						

ctt	ccc	gga	cac	gtc	ggc	ggc	gga	gag	ggg	ccc	acg	gct	gct	gcc	cg	240
Leu	Pro	Gly	His	Val	Gly	Gly	Glu	Gly	Gly	Pro	Thr	Ala	Ala	Ala	Arg	
			65				70			75		80				

cca	gag	act	cg	cg	cc	gag	cca	gct	ccc	cg	acc	cg	gcc	cc	gct	288
Pro	Glu	Thr	Arg	Arg	Pro	Glu	Pro	Ala	Pro	Arg	Thr	Arg	Ala	Pro	Ala	
			85				90			95						

ggc	aga	ccc	caa	ccc	agc	atg	agc	gct	gcc	acc	cac	tcg	ccc	atg	atg	336
Gly	Arg	Pro	Gln	Pro	Ser	Met	Ser	Ala	Ala	Thr	His	Ser	Pro	Met	Met	
			100				105			110						

cag	gtg	gc	tcc	ggc	aa	gg	gac	cg	gac	cct	ttg	cc	cc	gg	tgg	384
Gln	Val	Ala	Ser	Gly	Asn	Gly	Asp	Arg	Asp	Pro	Leu	Pro	Pro	Gly	Trp	
			115				120			125						

gag	atc	aag	atc	gac	cc	cag	acc	gg	ccc	ttc	ttc	gt	gac	cac	432	
Glu	Ile	Lys	Ile	Asp	Pro	Gln	Thr	Gly	Trp	Pro	Phe	Phe	Val	Asp	His	
			130				135			140						

aac	agc	cg	cc	act	ac	tg	aa	gac	cc	cg	cg	gt	cc	tct	gag	gg	480

Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly			
145	150	155	160
ccc aag gag act cca tcc tct gcc aat ggc cct tcc cgg gag ggc tct 528			
Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser			
165	170	175	
agg ctg ccg cct gct agg gaa ggc cac cct gtg tac ccc cag ctc cga 576			
Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg			
180	185	190	
cca ggc tac att ccc att cct gtg ctc cat gaa ggc gct gag aac cgg 624			
Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg			
195	200	205	
cag gtg cac cct ttc cat gtc tat ccc cag cct ggg atg cag cga ttc 672			
Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe			
210	215	220	
cga act gag gcg gca gca gcg gct cct cag agg tcc cag tca cct ctg 720			
Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu			
225	230	235	240
cgg ggc atg cca gaa acc act cag cca gat aaa cag tgt gga cag gtg 768			
Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val			
245	250	255	
gca gcg gcg gca gcc cag ccc cca gcc tcc cac gga cct gag cgg 816			
Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg			
260	265	270	
tcc cag tct cca gct gcc tct gac tgc tca tcc tca tcc tcc tcc gcc 864			
Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ala			
275	280	285	
agc ctg cct tcc tcc ggc agg agc agc ctg ggc agt cac cag ctc ccg 912			
Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro			
290	295	300	
cgg ggg tac atc tcc att ccg gtg ata cac gag cag aac gtt acc cgg 960			
Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg			
305	310	315	320
cca gca gcc cag ccc tcc ttc cac aaa gcc cag aag acg cac tac cca 1008			
Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro			
325	330	335	
gcg cag agg ggt gag tac cag acc cac cag cct gtg tac cac aag atc 1056			

Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile
 340 345 350 1104
 cag ggg gat gac tgg gag ccc cgg ccc ctg cgg gcg gca tcc ccc ttc
 Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe
 355 360 365 1152
 agg tca tct gtc cag ggt gca tcg agc cgg gag ggc tca cca gcc agg
 Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg
 370 375 380 1200
 agc agc acg cca ctc cac tcc ccc tcg ccc atc cgt gtg cac acc gtc
 Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val
 385 390 395 400 1248
 gtc gac agg cct cag cag ccc atg acc cat cga gaa act gca cct gtt
 Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val
 405 410 415 1296
 tcc cag cct gaa aac aaa cca gaa agt aag cca ggc cca gtt gga cca
 Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro
 420 425 430 1344
 gaa ctc cct gga cac atc cca att caa gtg atc cgc aaa gag gtc
 Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val
 435 440 445 1392
 gat tct aaa cct gtt tcc cag aag ccc cca cct ccc tct gag aag gta
 Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Ser Glu Lys Val
 450 455 460 1440
 gag gtg aaa gtt ccc cct gct cca gtt cct tgt cct cct ccc agc cct
 Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro
 465 470 475 480 1488
 ggc cct tct gct gtc ccc tct tcc ccc aag agt gtg gct aca gaa gag
 Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu
 485 490 495 1536
 agg gca gcc ccc agc act gcc cct gca gaa gct aca cct cca aaa cca
 Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro
 500 505 510 1584
 gga gaa gcc gag gct ccc cca aaa cat cca gga gtg ctg aaa gtg gaa
 Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu
 515 520 525 1632
 gcc atc ctg gag aag gtg cag ggg ctg gag cag gct gta gac aac ttt

Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe
 530 535 540

gaa ggc aag aag act gac aaa aag tac ctg atg atc gaa gag tat ttg 1680
 Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu
 545 550 555 560

acc aaa gag ctg ctg gcc ctg gat tca gtg gac ccc gag gga cga gcc 1728
 Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala
 565 570 575

gat gtg cgt cag gcc agg aga gac ggt gtc agg aag gtt cag acc atc 1776
 Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile
 580 585 590

ttg gaa aaa ctt gaa cag aaa gcc att gat gtc cca ggt caa gtc cag 1824
 Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln
 595 600 605

gtc tat gaa ctc cag ccc agc aac ctt gaa gca gat cag cca ctg cag 1872
 Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln
 610 615 620

gca atc atg gag atg ggt gcc gtg gca gca gac aag ggc aag aaa aat 1920
 Ala Ile Met Glu Met Gly Ala Val Ala Asp Lys Gly Lys Lys Asn
 625 630 635 640

gct gga aat gca gaa gat ccc cac aca gaa acc cag cca gaa gcc 1968
 Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala
 645 650 655

aca gca gca gcg act tca aac ccc agc agc atg aca gac acc cct ggt 2016
 Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly
 660 665 670

aac cca gca gca ccg tagcctctgc cctgtaaaag tcagactcgg aaccgatgtg 2071
 Asn Pro Ala Ala Pro
 675

tgcttaggg attttagttg catgcatttc agagacttta ggtcagttgg tttttagtt 2131

ctgcttggta tgcagtactt gggtgaggca aacactataa agggctaaaa gggaaaatga 2191

tgctttctt caatattctt actcttgtac aatataangaa gttgcttgg ttttggaaag 2251

tttaaccccg ttgcttggtc tgcagccctg tcnacttggg caccggacc acctgttagc 2311

tgggttggc cactgtctt tggtagctcg gactggaggg gtagatgggg agtcaattac 2371

ccatcacata aatatgaaac atttacaga aatgttgcctt ttttaatgag atgattttct 2431
tcatctcata attaaaatac ctgacttttag agagagtaaa atgtgccagg agccatagga 2491
atatctgtat gttggatgac tttaatgcta catttth 2528

<210> 6
<211> 677
<212> PRT
<213> Homo sapiens

<400> 6
Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Ala Asn Phe Ser Gly Leu
1 5 10 15
Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser
20 25 30
Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser
35 40 45
Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg
50 55 60
Leu Pro Gly His Val Gly Gly Glu Gly Pro Thr Ala Ala Ala Arg
65 70 75 80
Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala
85 90 95
Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met
100 105 110
Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp
115 120 125
Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His
130 135 140
Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly
145 150 155 160
Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser
165 170 175
Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg

180	185	190
Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg		
195	200	205
Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe		
210	215	220
Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu		
225	230	235
Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val		
245	250	255
Ala Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg		
260	265	270
Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ala		
275	280	285
Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro		
290	295	300
Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg		
305	310	315
Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro		
325	330	335
Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile		
340	345	350
Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe		
355	360	365
Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg		
370	375	380
Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val		
385	390	395
400		
Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val		
405	410	415
Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro		
420	425	430
Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val		

435

440

445

Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Pro Ser Glu Lys Val
450 455 460

Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro
465 470 475 480

Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu
485 490 495

Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro
500 505 510

Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu
515 520 525

Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe
530 535 540

Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu
545 550 555 560

Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala
565 570 575

Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile
580 585 590

Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln
595 600 605

Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln
610 615 620

Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn
625 630 635 640

Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala
645 650 655

Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly
660 665 670

Asn Pro Ala Ala Pro
675

<210> 7
<211> 1010
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (323)..(1009)

<400> 7
acgatatcct gtaagaccaa gaattgcaag gccagagttt gaattcttat acaaatggag 60
cgtatggtcc aacataccccc ccaggccctg gggcaaatac tgcctcatac tcaggggctt 120
attatgcacc tggttatact cagaccagtt actccacaga agttccaagt acttaccgtt 180
catctggcaa cagcccaact ccagtctctc gttggatcta tccccagcag gactgtcaag 240
actgaagcac cccctcttaa gggcaggtt ccaggatatc cgccctcaca gaaccctgga 300
atgaccctgc cccattatcc tt atg gag atg gta atc gta gtg ttc cac aat 352
Met Glu Met Val Ile Val Val Phe His Asn
1 5 10
cac ggc cga ctg tac gac cac aag aaa gat gcg tgg gct tct cct ggt 400
His Gly Arg Leu Tyr Asp His Lys Lys Asp Ala Trp Ala Ser Pro Gly
15 20 25
gct tat gga atg ggt ggc cgt tat ccc tgg cct tca tca gcg ccc tca 448
Ala Tyr Gly Met Gly Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser
30 35 40
gca cca ccc ggc aat ctc tac atg act gaa agt act tca cca tgg cct 496
Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro
45 50 55
ags agt ggc tct ccc cag tca ccc cct tca ccc cca gtc cag cag ccc 544
Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val Gln Gln Pro
60 65 70
aag gat tct tca tac ccc tat agc caa tca gat caa agc atg aac cgg 592
Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg
75 80 85 90
cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tcg ggg aca gtg 640
His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser Gly Thr Val
95 100 105

aac aat gat gat tca gat ctt ttg gat tcc caa gtc cag tat agt gct 688
 Asn Asn Asp Asp Ser Asp Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala
 110 115 120

gag cct cag ctg tat ggt aat gcc acc agt gac cat ccc aac aat caa 736
 Glu Pro Gln Leu Tyr Gly Asn Ala Thr Ser Asp His Pro Asn Asn Gln
 125 130 135

gat caa agt agc agt ctt cct gaa gaa tgg gta cct tca gat gaa agt 784
 Asp Gln Ser Ser Leu Pro Glu Glu Cys Val Pro Ser Asp Glu Ser
 140 145 150

act cct ccg agt att aaa aaa atc ata cat gtg ctg gag aag gtc cag 832
 Thr Pro Pro Ser Ile Lys Lys Ile Ile His Val Leu Glu Lys Val Gln
 155 160 165 170

tat ctt gaa caa gaa gta gaa gaa ttt gta gga aaa aag aca gac aaa 880
 Tyr Leu Glu Gln Glu Val Glu Phe Val Gly Lys Lys Thr Asp Lys
 175 180 185

gca tac tgg ctt ctg gaa gaa atg cta acc aag gaa ctt ttg gaa ctg 928
 Ala Tyr Trp Leu Leu Glu Glu Met Leu Thr Lys Glu Leu Leu Glu Leu
 190 195 200

gat tca gtt gaa act ggg ggc cag gac tct gta cgg cag gcc aga aaa 976
 Asp Ser Val Glu Thr Gly Gly Gln Asp Ser Val Arg Gln Ala Arg Lys
 205 210 215

gag gct gtt tgg aag att cag gcc ata ttg gaa a 1010
 Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu
 220 225

<210> 8
 <211> 229
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Glu Met Val Ile Val Val Phe His Asn His Gly Arg Leu Tyr Asp
 1 5 10 15

His Lys Lys Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly
 20 25 30

Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu
 35 40 45

Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln
50 55 60

Ser Pro Pro Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro
65 70 75 80

Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser
85 90 95

Val His Gln Tyr Glu Ser Ser Gly Thr Val Asn Asn Asp Asp Ser Asp
100 105 110

Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly
115 120 125

Asn Ala Thr Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu
130 135 140

Pro Glu Glu Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys
145 150 155 160

Lys Ile Ile His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val
165 170 175

Glu Glu Phe Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu
180 185 190

Glu Met Leu Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly
195 200 205

Gly Gln Asp Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile
210 215 220

Gln Ala Ile Leu Glu
225

<210> 9
<211> 689
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(482)

<220>
<221> unsure

<222> (105)

<223> any amino acid

<400> 9

ga gaa ata aaa aat gaa ctt ctc caa gca caa aac cct tct gaa ttg 47
 Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu
 1 5 10 15

tac ctg agc tcc aaa aca gaa ttg cag ggt tta att gga cag ttg gat 95
 Tyr Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp
 20 25 30

gag gta agt ntt gaa aaa aac ccc tgc atc cgg gaa gcc agg aga aga 143
 Glu Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg
 35 40 45

gca gtg atc gag gtg caa act ctg atc aca tat att gac ttg aag gag 191
 Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu
 50 55 60

gcc ctt gag aaa aga aag ctg ttt gct tgt gag gag cac cca tcc cat 239
 Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His
 65 70 75

aaa gcc gtc tgg aac gtc ctt gga aac ttg tct gag gag atc cag gga gaa 287
 Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu
 80 85 90 95

gtt ctt tca ttt gat gga aat cga acc gat aag aac tac atc cgg ctg 335
 Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu
 100 105 110

gaa gag ctg ctc acc aag cag ctg cta gcc ctg gat gct gtt gat ccg 383
 Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro
 115 120 125

cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt 431
 Gln Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu
 130 135 140

gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag 479
 Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu
 145 150 155

tac tgaaatacga gagatctcac ttttatact gttttgcact tcataatgtgc 532
 Tyr
 160

ttctatgtat agagagctt cagttcattt atttatacgt gcatatttca gtctcagtat 592
ttatgattga agcaaattct attcagtatc tgctgctttt gatgttgcaa gacaaatatc 652
attacagcac gttaactttt ccattcggat caaaaaa 689

<210> 10
<211> 160
<212> PRT
<213> Homo sapiens

<400> 10
Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr
1 5 10 15

Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu
20 25 30

Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg Ala
35 40 45

Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala
50 55 60

Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys
65 70 75 80

Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val
85 90 95

Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu
100 105 110

Glu Leu Leu Thr Lys Gln Leu Ala Leu Asp Ala Val Asp Pro Gln
115 120 125

Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala
130 135 140

Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu Tyr
145 150 155 160

<210> 11
<211> 246
<212> DNA
<213> Caenorhabditis elegans

<400> 11
atgtcttcc gcctcttcgt tgaaatattt cactttttt tccagctttt tccccatctc 60
gacctgctttt ggttttcga gaaaaccacg ttccaaatca gcgacatctc tcaaattgag 120
atcataggct ttttgaagat tgctcaaattt atgcttctca tattgcatga gcattttgaa 180
gcccgctca tcaacccaaag catttttcc acccatcaca atgattttat cattttctt 240
aaaatt 246

<210> 12
<211> 210
<212> PRT
<213> *Caenorhabditis elegans*

<400> 12
Met Lys Val Asn Val Ser Cys Ser Ser Val Gln Thr Thr Ile Asp Ile
1 5 10 15
Leu Glu Glu Asn Gln Gly Glu Asp Glu Ser Ile Leu Thr Leu Gly Gln
20 25 30
Leu Arg Asp Arg Ile Ala Thr Asp Asn Asp Val Asp Val Glu Thr Met
35 40 45
Lys Leu Leu His Arg Gly Lys Phe Leu Gln Gly Ala Asp Asp Val Ser
50 55 60
Leu Ser Thr Leu Asn Phe Lys Glu Asn Asp Lys Ile Ile Val Met Gly
65 70 75 80
Gly Lys Asn Ala Leu Val Asp Asp Ala Gly Phe Lys Met Leu Met Gln
85 90 95
Tyr Glu Lys His Asn Leu Ser Asn Leu Gln Lys Ala Tyr Asp Leu Asn
100 105 110
Leu Arg Asp Val Ala Asp Leu Glu Arg Gly Phe Leu Glu Lys Pro Lys
115 120 125
Gln Val Glu Met Gly Lys Lys Leu Glu Lys Lys Val Lys Tyr Phe Asn
130 135 140
Glu Glu Ala Glu Arg His Leu Glu Thr Leu Asp Gly Met Asn Ile Ile
145 150 155 160

Thr Glu Thr Thr Pro Glu Asn Gin Ala Lys Arg Asn Arg Glu Lys Arg
 165 170 175

Lys Thr Leu Val Asn Gly Ile Gln Thr Leu Leu Asn Gln Asn Asp Ala
 180 185 190

Leu Leu Arg Arg Leu Gln Glu Tyr Gln Ser Val Leu Asn Gly Asp Ile
 195 200 205

Pro Glu
 210

<210> 13

<211> 1377

<212> DNA

<213> *Caenorhabditis elegans*

<220>

<221> CDS

<222> (1)..(1377)

<400> 13

atg cca gtc gtg aac ata cca atc aaa ata ctt ggt cag aat caa tca 48
 Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser
 1 5 10 15

cat agt cga agt aac tcc tcg tct tct gtt gac aac gat cga aat caa 96
 His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
 20 25 30

cca cca cag cag cca cct caa ccg caa cca caa cag caa tct cag caa 144
 Pro Pro Gln Gln Pro Pro Gln Pro Gln Gln Gln Ser Gln Gln
 35 40 45

caa tac cag cag gct cca aac gtg aat acc aat atg cat cat tcc aac 192
 Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn
 50 55 60

gga ttc tca cct aac ttc cca tct cgt agt cct att ccg gac ttt ccc 240
 Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro
 65 70 75 80

agt ttt tca tct ggg ttc cca aac gat tct gaa tgg tct tcg aat ttc 288
 Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe
 85 90 95

ccg tcg ttt cca aat ttc cca agt gga ttc tca aat gga agt tct aat	336		
Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn			
100	105	110	
ttc cct gat ttt cca aga ttc gga aga gat gga gga cta tcg cca aac	384		
Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn			
115	120	125	
cca ccg atg caa gga tac agg aga agt cca aca cca aca tca act caa	432		
Pro Pro Met Gln Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gln			
130	135	140	
tct cca act tct aca tta aga cgc aac tct cag cag aat caa gct cct	480		
Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gln Ala Pro			
145	150	155	160
cca caa tat tct cag caa caa cca caa caa gct caa caa cgt cag aca	528		
Pro Gln Tyr Ser Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr			
165	170	175	
act cct ccg tca aca aaa gct tca tct cga cca cca tct cgt act cgt	576		
Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg			
180	185	190	
gaa cca aag gaa cct gag gta ccc gag aga cca gca gtt att cca ttg	624		
Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu			
195	200	205	
cca tat gag aag aag gag aaa cca ctg gag aag aaa ggt agt cgt gat	672		
Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp			
210	215	220	
tct gga aag ggt gat gag aac ctt gaa gag aac att gcc aag atc acg	720		
Ser Gly Lys Gly Asp Glu Asn Leu Glu Glu Asn Ile Ala Lys Ile Thr			
225	230	235	240
atc gga aag aat aat tgc gag tta tgt ccg gaa caa gaa acg gac ggc	768		
Ile Gly Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly			
245	250	255	
gac cca tct cca cta acc tcc cca atc acc gaa gga aag cca aag aga	816		
Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg			
260	265	270	
gga aag aaa ctt caa cgt aat caa agt gtt gtt gat ttc aat gcc aag	864		
Gly Lys Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys			
275	280	285	

aca att gtt act ttg gat aaa att gaa tta caa gtt gag cag ttg aga 912
 Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg
 290 295 300

 aaa aaa gct gct gaa ctc gaa atg gaa aaa gag caa att ctt cgt tct 960
 Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser
 305 310 315 320

 cta gga gaa atc agt gtt cat aac tgc atg ttc aaa ctg gaa gaa tgt 1008
 Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys
 325 330 335

 gat cgt gaa gag att gaa gca atc act gac cga ttg aca aaa aga aca 1056
 Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr
 340 345 350

 aag aca gtt caa gtt gtc gaa act cca cga aat gaa gaa cag aaa 1104
 Lys Thr Val Gln Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys
 355 360 365

 aaa gca ctg gaa gat gca act ttg atg atc gat gaa gtc gga gaa atg 1152
 Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met
 370 375 380

 atg cat tcg aat att gaa aag gct aag ctg tgc cta caa acc tac atg 1200
 Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met
 385 390 395 400

 aac gcc tgt tcg tac gaa gaa act gct gga gcc acc tgc caa aac ttc 1248
 Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe
 405 410 415

 ttg aag atc ata att cag tgc gct gct gat gat cag aaa cgc atc aag 1296
 Leu Lys Ile Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
 420 425 430

 cgt cgt ctg gaa aat ctg atg tct caa att gag aat gct gag aga acg 1344
 Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr
 435 440 445

 aaa gca gat ttg atg gat gat caa agc gaa tag 1377
 Lys Ala Asp Leu Met Asp Asp Gln Ser Glu
 450 455

<210> 14

<211> 458

<212> PRT

<213> *Caenorhabditis elegans*

<400> 14

Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser
1 5 10 15His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
20 25 30Pro Pro Gin Gln Pro Pro Gln Pro Gln Pro Gln Gln Ser Gin Gln
35 40 45Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn
50 55 60Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro
65 70 75 80Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe
85 90 95Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn
100 105 110Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn
115 120 125Pro Pro Met Gin Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gin
130 135 140Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gin Ala Pro
145 150 155 160Pro Gln Tyr Ser Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr
165 170 175Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg
180 185 190Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu
195 200 205Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp
210 215 220Ser Gly Lys Gly Asp Glu Asn Leu Glu Glu Asn Ile Ala Lys Ile Thr
225 230 235 240

Ile Gly Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly
245 250 255

Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg
260 265 270

Gly Lys Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys
275 280 285

Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg
290 295 300

Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser
305 310 315 320

Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys
325 330 335

Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr
340 345 350

Lys Thr Val Gln Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys
355 360 365

Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met
370 375 380

Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met
385 390 395 400

Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe
405 410 415

Leu Lys Ile Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
420 425 430

Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr
435 440 445

Lys Ala Asp Leu Met Asp Asp Gln Ser Glu
450 455

<210> 15
<211> 588
<212> DNA
<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1)...(588)

<400> 15

atg tca gaa aag act agc aca gtt aca ata cac tat gga aat cag cga 48
 Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg
 1 5 10 15

ttt ccg gta gca gtc aat cta aat gag acg tta agt gaa ctg att gat 96
 Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp
 20 25 30

gat tta ctt gaa acg act gag att tct gag aag aaa gtc aag ctt ttt 144
 Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe
 35 40 45

tac gct ggc aag cgt tta aaa gac aaa aaa gcc tcg tta tca aaa ttg 192
 Tyr Ala Gly Lys Arg Leu Lys Asp Lys Ala Ser Leu Ser Lys Leu
 50 55 60

ggc tta aaa aat cat agt aaa att cta tgt ata aga cca cat aag caa 240
 Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln
 65 70 75 80

caa cga ggt tcc aag gaa aaa gac acg gtt gag ccc gct ccg aaa gcg 288
 Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala
 85 90 95

gaa gcg gag aat cct gta ttt tcg cgt att tct gga gaa ata aaa gcc 336
 Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala
 100 105 110

atc gat cag tat gtt gac aaa gaa ctt tcc ccc atg tac gac aat tac 384
 Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr
 115 120 125

gta aat aaa ccg tcg aac gat cca aag cag aaa aac aaa cag aaa cta 432
 Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu
 130 135 140

atg ata agt gaa cta ctt tta caa cag ctt tta aaa ttg gat gga gtt 480
 Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val
 145 150 155 160

gac gta ctg ggc agc gag aaa ttg cgt ttt gaa cgg aag caa ctt gtt 528
 Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val
 165 170 175

tct aag atc caa aaa atg ttt gat cac gtt gac caa aca aqc caa gaa 576
Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu
180 185 190

gtg gcc gca tag 588
Val Ala Ala
195

<210> 16
<211> 195
<212> PRT
<213> Schizosaccharomyces pombe

<400> 16
Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg
1 5 10 15

Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp
20 25 30

Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe
35 40 45

Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu
50 55 60

Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln
65 70 75 80

Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala
85 90 95

Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala
100 105 110

Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr
115 120 125

Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu
130 135 140

Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val
145 150 155 160

Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val
165 170 175

Ser Lys Ile Gin Lys Met Leu Asp His Val Asp Gin Thr Ser Gin Giu
 180 185 190

Val Ala Ala
 195

<210> 17
 <211> 621
 <212> DNA
 <213> Schizosaccharomyces pombe

<220>
 <221> CDS
 <222> (1)...(621)

<400> 17
 atg tct ttt ttt acc cag ttg tgt tct atg gat aaa aaa tat tgg atc 48
 Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile
 1 5 10 15

tct cta gct gta ttg tca gtt act gtt ttg att agc gca tta ttg aaa 96
 Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys
 20 25 30

aag aga gct act gaa acc gaa gat att gtc gtt gtt cat tac gat ggc 144
 Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val Val His Tyr Asp Gly
 35 40 45

gaa aag ttg aat ttt gtg ttg cga caa cca agg ctg aat atg gtt tct 192
 Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser
 50 55 60

tac act agt ttt ctt cgt cgc gtg tgc aac gca ttt tca gta atg ccc 240
 Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro
 65 70 75 80

gac aaa gcg tct ctc aag tta aac ggg gtg acc ctc aag gat ggt tca 288
 Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser
 85 90 95

ctt tcc gac caa aat gtg caa aat gga agt gaa tta gag ctc gaa tta 336
 Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu Glu Leu
 100 105 110

ccc aaa ctg agc ccg gca atg caa caa att gaa gca tat ata gat gag 384
 Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu

115	120	125	
ctt caa cag gat ctc gtc cct aaa att gaa gcc ttc tgc caa tcg tct 432			
Leu Gln Gln Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser			
130	135	140	
ccc gct tcg gca caa gat gtt caa gat ttg cat aca cgc ctt agt gaa 480			
Pro Ala Ser Ala Gln Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu			
145	150	155	160
aca ttg ttg gct agg atg ata aaa tta gat gct gtt aat gtt gaa gac 528			
Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp			
165	170	175	
gac cca gaa gct cgt ctt aaa aga aaa gaa gct att cgt tta tct caa 576			
Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln			
180	185	190	
caa tat ttg agt aaa cta gat tcc acc aag aat caa aac aaa tga 621			
Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys			
195	200	205	
<210> 18			
<211> 206			
<212> PRT			
<213> Schizosaccharomyces pombe			
<400> 18			
Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile			
1	5	10	15
Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys			
20	25	30	
Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val Val His Tyr Asp Gly			
35	40	45	
Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser			
50	55	60	
Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro			
65	70	75	80
Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser			
85	90	95	
Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu Glu Leu			

100

105

110

Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu
115 120 125

Leu Gln Gin Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser
130 135 140

Pro Ala Ser Ala Gin Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu
145 150 155 160

Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp
165 170 175

Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln
180 185 190

Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys
195 200 205

<210> 19

<211> 2534

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (307)..(2034)

<400> 19

gcggagctcc gcatccaacc ccggggccgcg gccaacttct ctggactgga ccagaagttt 60

ctagccggcc agtgcttacc tccctttatc tcctccttcc cctctggcag cgaggaggct 120

atttccagac acttccaccc ctctctggcc acgtcacccc cgcccttaat tcataaaggt 180

gccccggcgcc ggcttcccggt acacgtcggc ggcggagagg ggcccacggc ggcggcccg 240

ccagagactc ggcccccggg gccagcggcc cgcacccgcg ccccaagcggg cagaccccaa 300

cccagc atg agc gcc acc cac tcg ccc atg atg cag gtg gcg tcc 348

Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser

1

5

10

ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg gag atc aag atc 396
Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile
15 20 25 30

gac ccg cag acc ggc tgg ccc ttc ttc gtg gac cac aac aac cgc acc 444
 Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr
 35 40 45

act acg tgg aac gac ccg cgc gtg ccc tct gag ggc ccc aag gag act 492
 Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr
 50 55 60

cca tcc tct gcc aat ggc cct tcc cgg gag ggc tct agg ctg ccg cct 540
 Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro
 65 70 75

gct agg gaa ggc cac cct gtg tac ccc cag ctc cga cca ggc tac att 588
 Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile
 80 85 90

ccc att cct gtg ctc cat gaa ggc gct gag aac ccg cag gtg cac cct 636
 Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg Gln Val His Pro
 95 100 110

ttc cat gtc tat ccc cag cct ggg atg cag cga ttc cga act gag gcg 684
 Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Glu Ala
 115 120 125

gca gca gcg gct cct cag agg tcc cag tca cct ctg cgg ggc atg cca 732
 Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro
 130 135 140

gaa acc act cag cca gat aaa cag tgt gga cag gtg gca gcg gcg gcg 780
 Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val Ala Ala Ala Ala
 145 150 155

gca gcc cag ccc cca gcc tcc cac gga cct gag cgg tcc cag tct cca 828
 Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg Ser Gln Ser Pro
 160 165 170

gct gcc tct gac tgc tca tcc tca tcc tcc tcg gcc agc ctg cct tcc 876
 Ala Ala Ser Asp Cys Ser Ser Ser Ser Ala Ser Leu Pro Ser
 175 180 185 190

tcc ggc agg agc agc ctg ggc agt cac cag ctc ccg cgg ggg tac atc 924
 Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro Arg Gly Tyr Ile
 195 200 205

tcc att ccg gtg ata cac gag cag aac gtt acc cgg cca gca gcc cag 972
 Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg Pro Ala Ala Gln
 210 215 220

ccc tcc ttc cac aaa gcc cag aag acg cac tac cca gcg cag agg ggt 1020
 Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro Ala Gln Arg Gly
 225 230 235

gag tac cag acc cac cag cct gtg tac cac aag atc cag ggg gat gac 1068
 Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile Gln Gly Asp Asp
 240 245 250

tgg gag ccc cgg ccc ctg cgg gcg gca tcc ccg ttc agg tca tct gtc 1116
 Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe Arg Ser Ser Val
 255 260 265 270

cag ggt gca tcg agc cgg gag ggc tca cca gcc agg agc agc acg cca 1164
 Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg Ser Ser Thr Pro
 275 280 285

ctc cac tcc ccc tcg ccc atc cgt gtg cac acc gtg gtc gac agg cct 1212
 Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val Val Asp Arg Pro
 290 295 300

cag cag ccc atg acc cat cga gaa act gca cct gtt tcc cag cct gaa 1260
 Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val Ser Gln Pro Glu
 305 310 315

aac aaa cca gaa agt aag cca ggc cca gtt gga cca gaa ctc cct cct 1308
 Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro Glu Leu Pro Pro
 320 325 330

gga cac atc cca att caa gtg atc cgc aaa gag gtg gat tct aaa cct 1356
 Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val Asp Ser Lys Pro
 335 340 345 350

gtt tcc cag aag ccc cca cct ccc tct gag aag gta gag gtg aaa gtt 1404
 Val Ser Gln Lys Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val
 355 360 365

ccc cct gct cca gtt cct tgt cct ccc agc cct ggc cct tct gct 1452
 Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro Gly Pro Ser Ala
 370 375 380

gtc ccc tct tcc ccc aag agt gtg gct aca gaa gag agg gca gcc ccc 1500
 Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro
 385 390 395

agc act gcc cct gca gaa gct aca cct cca aaa cca gga gaa gcc gag 1548
 Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu
 400 405 410

gct ccc cca aaa cat cca gga gtg ctg aaa gtg gaa gcc atc ctg gag 1596
 Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu
 415 420 425 430

aag gtg cag ggg ctg gag cag gct gta gac aac ttt gaa ggc aag aag 1644
 Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys
 435 440 445

act gac aaa aag tac ctg atg atc gaa gag tat ttg acc aaa gag ctg 1692
 Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu
 450 455 460

ctg gcc ctg gat tca gtg gac ccc gag gga cga gcc gat gtg cgt cag 1740
 Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln
 465 470 475

gcc agg aga gac ggt gtc agg aag gtt cag acc atc ttg gaa aaa ctt 1788
 Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu
 480 485 490

gaa cag aaa gcc att gat gtc cca ggt caa gtc cag gtc tat gaa ctc 1836
 Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu
 495 500 505 510

cag ccc agc aac ctt gaa gca gat cag cca ctg cag gca atc atg gag 1884
 Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu
 515 520 525

atg ggt gcc gtg gca gca gac aag ggc aag aaa aat gct gga aat gca 1932
 Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala
 530 535 540

gaa gat ccc cac aca gaa acc cag cag cca gaa gcc aca gca gca gcg 1980
 Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala
 545 550 555

act tca aac ccc agc agc atg aca gac acc cct ggt aac cca gca gca 2028
 Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala
 560 565 570

ccg tag cctctgccct gtaaaaatca gactcgaaac cgatgtgtgc ttttaggaaat 2084
 Pro
 575

tttaagttgc atgcatttca gagactttaa gtcagttgggt ttttatttagc tgcttggat 2144

gcagtaactt gggtggaggc aaaacactaa taaaagggtc aaaaaggaaa atgatgcttt 2204

tcttctataat tcttactctg tacaaataaaa gaagtttgtt 2264
 ccgttgccttgc ttcgtgcagcc ctgtctactt gggcacccccc accacactgtt agctgtggtt 2324
 gtgcactgtc tttttagtct ctggactgga ggggttagatg gggagtcata tacccatcac 2384
 ataaatatga aacatttatac agaaatgttg ccatttaat gagatgattt tcttcatctc 2444
 ataattaaaa tacctgactt tagagagagt aaaatgtgcc aggagccata ggaatatctg 2504
 tatgttggat gactttaatg ctacatttc 2534

<210> 20
 <211> 575
 <212> PRT
 <213> Homo sapiens

<400> 20
 Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser Gly Asn
 1 5 10 15
 Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile Asp Pro
 20 25 30
 Gin Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr Thr Thr
 35 40 45
 Trp Asn Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr Pro Ser
 50 55 60
 Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro Ala Arg
 65 70 75 80
 Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile Pro Ile
 85 90 95
 Pro Val Leu His Glu Gly Ala Glu Asn Arg Gin Val His Pro Phe His
 100 105 110
 Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Glu Ala Ala Ala
 115 120 125
 Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro Glu Thr
 130 135 140
 Thr Gln Pro Asp Lys Gin Cys Gly Gin Val Ala Ala Ala Ala Ala

145	150	155	160
Gln Pro Pro Ala Ser His Gly Pro Glu Arg Ser Gln Ser Pro Ala Ala			
165	170	175	
Ser Asp Cys Ser Ser Ser Ser Ala Ser Leu Pro Ser Ser Gly			
180	185	190	
Arg Ser Ser Leu Gly Ser His Gln Leu Pro Arg Gly Tyr Ile Ser Ile			
195	200	205	
Pro Val Ile His Glu Gln Asn Val Thr Arg Pro Ala Ala Gln Pro Ser			
210	215	220	
Phe His Lys Ala Gln Lys Thr His Tyr Pro Ala Gln Arg Gly Glu Tyr			
225	230	235	240
Gln Thr His Gln Pro Val Tyr His Lys Ile Gln Gly Asp Asp Trp Glu			
245	250	255	
Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe Arg Ser Ser Val Gln Gly			
260	265	270	
Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg Ser Ser Thr Pro Leu His			
275	280	285	
Ser Pro Ser Pro Ile Arg Val His Thr Val Val Asp Arg Pro Gln Gln			
290	295	300	
Pro Met Thr His Arg Glu Thr Ala Pro Val Ser Gln Pro Glu Asn Lys			
305	310	315	320
Pro Glu Ser Lys Pro Gly Pro Val Gly Pro Glu Leu Pro Pro Gly His			
325	330	335	
Ile Pro Ile Gln Val Ile Arg Lys Glu Val Asp Ser Lys Pro Val Ser			
340	345	350	
Gln Lys Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val Pro Pro			
355	360	365	
Ala Pro Val Pro Cys Pro Pro Pro Ser Pro Gly Pro Ser Ala Val Pro			
370	375	380	
Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro Ser Thr			
385	390	395	400
Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu Ala Pro			

405

410

415

Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val
 420 425 430

Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp
 435 440 445

Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala
 450 455 460

Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg
 465 470 475 480

Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln
 485 490 495

Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro
 500 505 510

Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly
 515 520 525

Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp
 530 535 540

Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser
 545 550 555 560

Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala Pro
 565 570 575

<210> 21

<211> 1966

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (43)..(1416)

<400> 21

cggtgggagc gggcgggaa gcgcttcagg gcaagcggatc cc atg tcg gcc ctg 54
 Met Ser Ala Leu
 1

agg cgc tcg ggc tac ggc ccc agt gac ggt ccg tcc tac ggc cgc tac 102

Arg Arg Ser Gly Tyr Gly Pro Ser Asp Gly Pro Ser Tyr Gly Arg Tyr			
5	10	15	20
tac ggg cct ggg ggt gga gat gtg ccg gta cac cca cct cca ccc tta			150
Tyr Gly Pro Gly Gly Asp Val Pro Val His Pro Pro Pro Pro Leu			
25	30	35	
tat cct ctt cgc cct gaa cct ccc cag cct ccc att tcc tgg cgg gtg			198
Tyr Pro Leu Arg Pro Glu Pro Pro Gln Pro Pro Pro Ile Ser Trp Arg Val			
40	45	50	
cgc ggg ggc ggc ccg gcg gag acc acc tgg ctg gga gaa ggc gga gga			246
Arg Gly Gly Pro Ala Glu Thr Thr Trp Leu Gly Glu Gly Gly			
55	60	65	
ggc gat ggc tac tat ccc tcg gga ggc gcc tgg cca gag cct ggt cga			294
Gly Asp Gly Tyr Tyr Pro Ser Gly Gly Ala Trp Pro Glu Pro Gly Arg			
70	75	80	
gcc gga gga agc cac cag gag cag cca cca tat cct agc tac aat tct			342
Ala Gly Gly Ser His Gln Glu Gln Pro Pro Tyr Pro Ser Tyr Asn Ser			
85	90	95	100
aac tat tgg aat tct act gcg aga tct agg gct cct tac cca agt aca			390
Asn Tyr Trp Asn Ser Thr Ala Arg Ser Arg Ala Pro Tyr Pro Ser Thr			
105	110	115	
tat cct gta aga cca gaa ttg caa ggc cag agt ttg aat tct tat aca			438
Tyr Pro Val Arg Pro Glu Leu Gln Gly Gln Ser Leu Asn Ser Tyr Thr			
120	125	130	
aat gga gcg tat ggt cca aca tac ccc cca ggc cct ggg gca aat act			486
Asn Gly Ala Tyr Gly Pro Thr Tyr Pro Pro Gly Pro Gly Ala Asn Thr			
135	140	145	
gcc tca tac tca ggg gct tat tat gca cct ggt tat act cag acc agt			534
Ala Ser Tyr Ser Gly Ala Tyr Tyr Ala Pro Gly Tyr Thr Gln Thr Ser			
150	155	160	
tac tcc aca gaa gtt cca agt act tac cgt tca tct ggc aac agc cca			582
Tyr Ser Thr Glu Val Pro Ser Thr Tyr Arg Ser Ser Gly Asn Ser Pro			
165	170	175	180
act cca gtc tct cgt tgg atc tat ccc cag cag gac tgt cag act gaa			630
Thr Pro Val Ser Arg Trp Ile Tyr Pro Gln Gln Asp Cys Gln Thr Glu			
185	190	195	
gca ccc cct ctt agg ggg cag gtt cca gga tat ccg cct tca cag aac			678

Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro Pro Ser Gln Asn			
200	205	210	
cct gga atg acc ctg ccc cat tat cct tat gga gat ggt aat cgt agt			726
Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp Gly Asn Arg Ser			
215	220	225	
gtt cca caa tca gga ccg act gta cga cca caa gaa gat gcg tgg gct			774
Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu Asp Ala Trp Ala			
230	235	240	
tct cct ggt gct tat gga atg ggt ggc cgt tat ccc tgg cct tca tca			822
Ser Pro Gly Ala Tyr Gly Met Gly Arg Tyr Pro Trp Pro Ser Ser			
245	250	255	260
gcg ccc tca gca cca ccc ggc aat ctc tac atg act gaa agt act tca			870
Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser			
265	270	275	
cca tgg cct agc agt ggc tct ccc cag tca ccc cct tca ccc cca gtc			918
Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val			
280	285	290	
cag cag ccc aag gat tct tca tac ccc tat agc caa tca gat caa agc			966
Gln Gln Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser			
295	300	305	
atg aac cgg cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tcg			1014
Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser			
310	315	320	
ggg aca gtg atc aat gaa gat tca gat ctt ttg gat tcc caa gtc cag			1062
Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp Ser Gln Val Gln			
325	330	335	340
tat agt gct gag cct cag ctg tat ggt aat gcc acc agt gac cat ccc			1110
Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr Ser Asp His Pro			
345	350	355	
aac aat caa gat caa agt agc agt ctt cct gaa gaa tgt gta cct tca			1158
Asn Asn Gln Asp Gln Ser Ser Ser Leu Pro Glu Glu Cys Val Pro Ser			
360	365	370	
gat gaa agt act cct ccg agt att aaa aaa atc ata cat gtg ctg gag			1206
Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile His Val Leu Glu			
375	380	385	
aag gtc cag tat ctt gaa caa gaa gta gaa gaa ttt gta gga aaa aag			1254

Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe Val Gly Lys Lys
 390 395 400

aca gac aaa gca tac tgg ctt ctg gaa gaa atg cta acc aag gaa ctt 1300
 Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu Thr Lys Glu Leu
 405 410 415 420

ttg gaa ctg gat tca gtt gaa act ggg ggc cag gac tct gta cgg cag 1350
 Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp Ser Val Arg Gln
 425 430 435

gcc aga aaa gag gct gtt tgt aag att cag gcc ata ctg gaa aaa tta 1398
 Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu Lys Leu
 440 445 450

gaa aaa aaa gga tta tga aaggatttag aacaaagtgg aagcctgtta 1446
 Glu Lys Lys Gly Leu
 455

ctaacttgac caaagaacac ttgatttagt taattaccct cttttgaaa tgcctgtga 1506
 tgacaagaag caatacattc cagtttcc tttgattta tacttgaaaa actggcaaag 1566

gaatggaaga atattttagt catgaagttg tttcagttt tcagacgaat gaatgtaata 1626

ggaaactatg gagttaccaa tattgccaag tagactcaact ccttaaaaaa tttatggata 1686

tctacaagct gcttattacc agcaggaggg aaacacactt cacacaacag gcttatcaga 1746

aacctaccag atgaaactgg atataatttg agacaaacag gatgtgtttt tttaaacatc 1806

tggatatctr gtcacatttt tgcacattgt gactgcttc aacatataact tcacgtgtaa 1866

ttatagctt gacttagcc ttcttgact tctgtttgt tttgttattt gcagtttaca 1926

aatatagtat tattctctaa aaaaaaaaaa aaaaaaaaaa 1966

<210> 22
 <211> 457
 <212> PRT
 <213> Homo sapiens

<400> 22
 Met Ser Ala Leu Arg Arg Ser Gly Tyr Gly Pro Ser Asp Gly Pro Ser
 1 5 10 15

Tyr Gly Arg Tyr Tyr Gly Pro Gly Gly Asp Val Pro Val His Pro

20

25

30

Pro Pro Pro Leu Tyr Pro Leu Arg Pro Glu Pro Pro Gln Pro Pro Ile
35 40 45

Ser Trp Arg Val Arg Gly Gly Gly Pro Ala Glu Thr Thr Trp Leu Gly
50 55 60

Glu Gly Gly Gly Asp Gly Tyr Tyr Pro Ser Gly Gly Ala Trp Pro
65 70 75 80

Glu Pro Gly Arg Ala Gly Gly Ser His Gln Glu Gln Pro Pro Tyr Pro
85 90 95

Ser Tyr Asn Ser Asn Tyr Trp Asn Ser Thr Ala Arg Ser Arg Ala Pro
100 105 110

Tyr Pro Ser Thr Tyr Pro Val Arg Pro Glu Leu Gln Gly Gln Ser Leu
115 120 125

Asn Ser Tyr Thr Asn Gly Ala Tyr Gly Pro Thr Tyr Pro Pro Gly Pro
130 135 140

Gly Ala Asn Thr Ala Ser Tyr Ser Gly Ala Tyr Tyr Ala Pro Gly Tyr
145 150 155 160

Thr Gln Thr Ser Tyr Ser Thr Glu Val Pro Ser Thr Tyr Arg Ser Ser
165 170 175

Gly Asn Ser Pro Thr Pro Val Ser Arg Trp Ile Tyr Pro Gln Gln Asp
180 185 190

Cys Gln Thr Glu Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro
195 200 205

Pro Ser Gln Asn Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp
210 215 220

Gly Asn Arg Ser Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu
225 230 235 240

Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Arg Tyr Pro
245 250 255

Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr
260 265 270

Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro

275	280	285
Ser Pro Pro Val Gin Gin Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln		
290	295	300
Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln		
305	310	315
Tyr Glu Ser Ser Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp		
325	330	335
Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr		
340	345	350
Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu Pro Glu Glu		
355	360	365
Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile		
370	375	380
His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe		
385	390	395
Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu		
405	410	415
Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp		
420	425	430
Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile		
435	440	445
Leu Glu Lys Leu Glu Lys Lys Gly Leu		
450	455	

<210> 23
<211> 4308
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (247)..(1590)

<400> 23
cccccccccc cccccccccc ccngaagacg cccggagcgg ctgctgcagc cagtagcggc 60

cccttcacccg gctgccccgc tcagacctag tcgggagggg tgcgaggcat gcaactgggg 120
 gcccagctcc ggtgccgcac cccgtaaagg gctgatcttc cacctcgcca cctcagccac 180
 gggacgccaa gaccgcatcc aattcagact tctttggtg cttgtgaaac tgaacacaac 240
 aaaagt atg gat atg gga aac caa cat cct tct att agt agg ctt cag 288
 Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln
 1 5 10

gaa atc caa aag gaa gta aaa agt gta gaa cag caa gtt atc ggc ttc 336
 Glu Ile Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe
 15 20 25 30

agt ggt ctg tca gat gac aag aat tac aag aaa ctg gag agg att cta 384
 Ser Gly Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu
 35 40 45

aca aaa cag ctt ttt gaa ata gac tct gta gat act gaa gga aaa gga 432
 Thr Lys Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly
 50 55 60

gat att cag caa gct agg aag cgg gca gca cag gag aca gaa cgt ctt 480
 Asp Ile Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu
 65 70 75

ctc aaa gag ttg gag cag aat gca aac cac cca cac cgg att gaa ata 528
 Leu Lys Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile
 80 85 90

cag aac att ttt gag gaa gcc cag tcc ctc gtg aga gag aaa att gtg 576
 Gln Asn Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val
 95 100 105 110

cca ttt tat aat gga ggc aac tgc gta act gat gag ttt gaa gaa ggc 624
 Pro Phe Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly
 115 120 125

atc caa gat atc att ctg agg ctg aca cat gtt aaa act gga gga aaa 672
 Ile Gln Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys
 130 135 140

atc tcc ttg cgg aaa gca agg tat cac act tta acc aaa atc tgt gcg 720
 Ile Ser Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala
 145 150 155

gtg caa gag ata atc gaa gac tgc atg aaa aag cag cct tcc ctg ccg 768
 Val Gln Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro

160	165	170	
ctt tcc gag gat gca cat cct tcc gtt gcc aaa atc aac ttc gtg atg 816			
Leu Ser Glu Asp Ala His Pro Ser Val Ala Lys Ile Asn Phe Val Met			
175	180	185	190
tgc gag gtg aac aag gcc cga ggg gtc ctg att gca ctt ctg atg ggt 864			
Cys Glu Val Asn Lys Ala Arg Gly Val Leu Ile Ala Leu Leu Met Gly			
195	200	205	
gtg aac aac aat gag acc tgc agg cac tta tcc tgt gtg ctc tcg ggg 912			
Val Asn Asn Asn Glu Thr Cys Arg His Leu Ser Cys Val Leu Ser Gly			
210	215	220	
ctg atc gct gac ctg gat gct cta gat gtg tgc ggc cgg aca gaa atc 960			
Leu Ile Ala Asp Leu Asp Ala Leu Asp Val Cys Gly Arg Thr Glu Ile			
225	230	235	
aga aat tat cgg agg gag gta gta gaa gat atc aac aaa tta ttg aaa 1008			
Arg Asn Tyr Arg Arg Glu Val Val Glu Asp Ile Asn Lys Leu Leu Lys			
240	245	250	
tat ctg gat ttg gaa gag gaa gca gac aca act aaa gca ttt gac ctg 1056			
Tyr Leu Asp Leu Glu Glu Ala Asp Thr Thr Lys Ala Phe Asp Leu			
255	260	265	270
aga cag aat cat tcc att tta aaa ata gaa aag gtc ctc aag aga atg 1104			
Arg Gln Asn His Ser Ile Leu Lys Ile Glu Lys Val Leu Lys Arg Met			
275	280	285	
aga gaa ata aaa aat gaa ctt ctc caa gca caa aac cct tct gaa ttg 1152			
Arg Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu			
290	295	300	
tac ctg agc tcc aaa aca gaa ttg cag ggt tta att gga cag ttg gat 1200			
Tyr Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp			
305	310	315	
gag gta agt ctt gaa aaa aac ccc tgc atc cgg gaa gcc agg aga aga 1248			
Glu Val Ser Leu Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg			
320	325	330	
gca gtg atc gag gtg caa act ctg atc aca tat att gac ttg aag gag 1296			
Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu			
335	340	345	350
gcc ctt gag aaa aga aag ctg ttt gct tgt gag gag cac cca tcc cat 1344			
Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His			

355

360

365

aaa gcc gtc tgg aac gtc ctt gga aac ttg tct gag atc cag gga gaa 1392
 Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu
 370 375 380

gtt ctt tca ttt gat gga aat cga acc gat aag aac tac atc cgg ctg 1440
 Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu
 385 390 395

gaa gag ctg ctc acc aag cag ctg cta gcc ctg gat gct gtt gat ccg 1488
 Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro
 400 405 410

cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt 1536
 Gln Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu
 415 420 425 430

gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag 1584
 Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu
 435 440 445

tac tga aataccagag atctcacttt tgatactgtt ttgcacttca tatgtgcttc 1640
 Tyr

tatgtataga gagctttcag ttcattgatt tatacgtgca tatttcagtc tcagtattta 1700

tgattgaagc aaattctatt cagtatctgc tgcgtttgat gttgcaagac aaatatcatt 1760

acagcacgtt aactttcca ttcggatcat tatctgtatg atgtgggttg gtttgggtgg 1820

tttgccttt ttttgcgtt ttaatcaga aaacaaaata gaggcagctt ttgttagattt 1880

taaatgggtt gtgcaagcat taaaatgcag gtcttcaga atctagaact aggccataacc 1940

ttacataata ctagaaaaat tatgagaaaag gggaaatttt tggttaaata agagtaagg 2000

tcaaacacaa gcagtacatg ttctgtttca ttatgtcaga tagaaggctt tttttcact 2060

tataaggcct gattggcctt acccagctt acgggggtggg gttttttgtt ttgttcagac 2120

agtctgttct tttgtaaaca ttttagttg gaaaaacagc atctgcattt tccccatcct 2180

ctacgtttta gagaggaatc ttgtttttgt gtgcaacata agaaaattat gaaaactaat 2240

agccaaaaaaa cctttgagat tgcattaaag agaagggata aaggaccagc aataatacct 2300

tgttaagttgc ttttgggttgaaaatctgag cttatagttt tccttagtga gtaaattcat 2360

aaggatggga acattnaat taagttaatg ggcctttaaa aaaaaaaaaaag gaaacactca 2420
tacctgttgtgaggatga atactggaga cgggttacca atgtcaggtt atactaaaac 2480
taaatcagaa agtctgaatg tagcacataa tggttctctt ctgttgcctt aggctgtaaa 2540
atggacagcc ttgtcacacc tccccgggtgc tggttacaa cgtgagggtt gacgctgtca 2600
gtaaccaga gggaccaggc cttccttaggt tttcttaggca gtcagctgtt aaccactcac 2660
ttagtaatg tcataactac acctgctcca ggaccaatca gtgaaacctg ctcggattta 2720
aaggcttcct ctgggtgcct gctgaacaac tgagctcatg tcatggcat gtggtggttt 2780
ctctgttgcc tgaagagagcc attaaagtca gtcgtgcgtg aagcatctct cttctaaagg 2840
atgtgtatccatccataatgc ttcttgagga tccggtacaa aatgatttcc caaagttctg 2900
aagtgccttg agaacatgtg ggtccgagtg ttataacaga ctcctcccccc gggcacctt 2960
ttgcctggtc atcctgttag agtacatctt tggaaatcca gggtaatatt ctctttcaga 3020
gatgctcatt gtgttaactct gtgttagggag atagtcactt taaacagctc aaagtagctt 3080
gctaaaggag tagccttaaa tacctaaaag atgacagaag catagccctt aacaaatctt 3140
cagcttgcctt ctcagtatccatccatg aaaatccctt gctatgtctt tcctactaga 3200
aatgttcttag aatcgctgga cggtggttc agagggcagt cggtatttag gccgtgagct 3260
tcccatacta ctgcagggtcc aactcctggc aaccgcgggc tcaaggcagg tcattggat 3320
ccacgttttgcacagtag ttgttaggatt gctttctgt atcataatccatccatg 3380
ttaaaatctt gaggaagagt ttttattttt tatttatttt tgagatggag tctctgttgc 3440
ccaggctgca gtgcagtggc gccatctcag ctcactgcaa cctccacccctt ccaggttcaa 3500
gcgattctcc tgcctcagcc acctgagtag ctgggagttac aggcattgtgg caccatgcctt 3560
ggctaatttt ttttttttta atagagtttta gatttcacca tttatggcag gctggtctcg 3620
aactcctgac ctcgtgatcc gcccgcctcg gccccccaaa gtgctggat taacgggtgt 3680
gagccacggc gcccagccca ggaagagttt ttaaattttaga gctctgttta attataaccac 3740
tggaaatca tggttacgct tcaggcatat tcttccccag agtactactt acattttaaa 3800

tttcattttg taaagttaaa tgcgcattt ccctttaaaa gtgtccattt ttctttgaaa 3860
 gtagacgttt cagtcattct tttcaaaacaa gtgttgcgtt acctttgcc aagctgtggg 3920
 catcggtgtt gagtacaggg tgctcagctc ttccaccgtc attttgaatt gttcacatgg 3980
 gtaattggtc atggaaatga tcagattgac cttgattgac tgtcaggcat ggctttgtt 4040
 ctatgttcaa tctgttctcg ttccattgtac cggattattc tactcctgca atgaaccctg 4100
 ttgacacccgg atttagctct tgcggcctt cgtggggagc tggttgcgtt aatatgagct 4160
 actgcatgta attcttaaac tgggcttgc acattgtatt gtattttgt gatctgtaat 4220
 gaaaagaatc tgtactgcaa gtaaaaccta ctccccaaaa atgtgtggct ttgggtctgc 4280
 attaaacgct gtagtccatg ttcatgcc 4308

<210> 24
 <211> 447
 <212> PRT
 <213> Homo sapiens

<400> 24
 Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gin Glu Ile
 1 5 10 15
 Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe Ser Gly
 20 25 30
 Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu Thr Lys
 35 40 45
 Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly Asp Ile
 50 55 60
 Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu Leu Lys
 65 70 75 80
 Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile Gln Asn
 85 90 95
 Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val Pro Phe
 100 105 110
 Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Gly Ile Gln

115	120	125
Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys Ile Ser		
130	135	140
Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala Val Gln		
145	150	155
160		
Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro Leu Ser		
165	170	175
Glu Asp Ala His Pro Ser Val Ala Lys Ile Asn Phe Val Met Cys Glu		
180	185	190
Val Asn Lys Ala Arg Gly Val Leu Ile Ala Leu Leu Met Gly Val Asn		
195	200	205
Asn Asn Glu Thr Cys Arg His Leu Ser Cys Val Leu Ser Gly Leu Ile		
210	215	220
Ala Asp Leu Asp Ala Leu Asp Val Cys Gly Arg Thr Glu Ile Arg Asn		
225	230	235
240		
Tyr Arg Arg Glu Val Val Glu Asp Ile Asn Lys Leu Leu Lys Tyr Leu		
245	250	255
Asp Leu Glu Glu Ala Asp Thr Thr Lys Ala Phe Asp Leu Arg Gln		
260	265	270
Asn His Ser Ile Leu Lys Ile Glu Lys Val Leu Lys Arg Met Arg Glu		
275	280	285
Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr Leu		
290	295	300
Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu Val		
305	310	315
320		
Ser Leu Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Ala Val		
325	330	335
Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala Leu		
340	345	350
Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys Ala		
355	360	365
Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val Leu		

370

375

380

Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu Glu
385 390 395 400

Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln Gly
405 410 415

Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala Gln
420 425 430

Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu Tyr
435 440 445

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21053

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :07N 21/02; C07K 1/00

US CL :530/387.1, 350; 435/6, 7/1; 536/23.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/387.1, 350; 435/6, 7/1; 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,652,223 A (KOHN ET AL) 29 July 1997(29/7/97) see entire document.	2-5, 14, 32-34
X	Database Genbank-EST, National Center for Biotech. Info., Accession No. AA693697, HILLIER, L. ET AL. 'WashU-NCI human EST Project,' 16 December 1997, see entire reference.	2
X	Database Genbank-EST, National Center for Biotech. Info., Accession No. AA456862, NCI_CGAP, 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index,' 15 August 1997, see entire reference.	2,4

 Further documents are listed in the continuation of Box C. See patent family annex.

A	Special categories of cited documents:	*T*	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
B	document defining the general state of the art which is not considered to be of particular relevance	*X*	document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
C	earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
D	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z*	document member of the same patent family
E	document referring to an oral disclosure, use, exhibition or other means		
F	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

24 NOVEMBER 1999

Date of mailing of the international search report

19 JAN 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized Officer

SHEELA J. HUFF

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No

PCT/US99/21053

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: 1, 13, 24, 25 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

No meaningful search could be carried out because no limitations could be placed on the sequence

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest

No protest accompanied the payment of additional search fees

INTERNATIONAL SEARCH REPORT

International application No
PCT/US99/21053

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genbank, National Center for Biotech. Info., Accession No. G29287, MYERS, R.M., 04 October 1996, see entire reference.	2,4
X	Database Genbank, National Center for Biotech. Info., Accession No. G06974, HUDSON, T., "Whitehead Institute/MIT Center for Genome Research,'19 October 1995, see entire reference.	2,4
X	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. V81267, OTSUKA PHARM CO LTD, 'New Bcl-2 interaction protein gene (Bis)- useful for elucidation of the molecular mechanism of apoptosis, and in diagnosis, prevention and treatment of diseases,' 15 December 1998 see entire reference.	2-5
X	Database, Geneseq, Derwent, Alexandria, Virginia, Accession No. T19051, MATSUBARA ET AL., "Identifying gene signatures in 3'-directed human cDNA library,' 01 June 1995, see entire reference.	2,4
X	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. Q90296, LA JOLLA CANCER RES FOUN. 'Human Bcl-2-associated protein BAG-1 cDNA,'18 May 1995 see entire reference.	2-5,14



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :	A1	(11) International Publication Number: WO 00/14106
C07K 1/00		(43) International Publication Date: 16 March 2000 (16.03.00)

(21) International Application Number: PCT/US99/21053	(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(22) International Filing Date: 9 September 1999 (09.09.99)	
(30) Priority Data: 09/150,489 9 September 1998 (09.09.98) US	Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(71) Applicant: THE BURNHAM INSTITUTE [US/US]; 10901 N. Torrey Pines Road, La Jolla, CA 92037 (US).	
(72) Inventors: REED, John, C.; 17044 El Camino Real, Rancho Santa Fe, CA 92067 (US). TAKAYAMA, Shinichi; 390 Stratford Court #3, Del Mar, CA 92014 (US).	
(74) Agents: WONG, James, J. et al.; Campbell & Flores LLP, Suite 700, 4370 La Jolla Village Drive, San Diego, CA 92122 (US).	

(54) Title: NOVEL BAG PROTEINS AND NUCLEIC ACID MOLECULES ENCODING THEM

(57) Abstract

The present invention provides a family of BAG-1 related proteins from humans (BAG-1L, BAG-1, BAG-2, BAG-3, BAG-4 and BAG-5), the invertebrate *C. elegans* (BAG-1, BAG-2) and the fission yeast *S. pombe* (BAG-1A, BAG-1B) and the nucleic acid molecules that encode them.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL BAG PROTEINS AND
NUCLEIC ACID MOLECULES ENCODING THEM

STATEMENT AS TO RIGHTS TO INVENTIONS MADE
UNDER FEDERALLY-SPONSORED RESEARCH AND DEVELOPMENT

5 This invention was made with government support under grant number CA-67329 awarded by the National Institutes of Health. The United States Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

10

FIELD OF THE INVENTION

This invention relates generally to the fields of molecular biology and molecular medicine and more specifically to a novel family of proteins that can regulate protein folding. The functions of these proteins 15 are potentially diverse, including promoting tumor cell growth and metastasis.

BACKGROUND INFORMATION

The Hsc70/Hsp70-family of molecular chaperones participate in protein folding reactions, controlling 20 protein bioactivity, degradation, complex assembly/disassembly, and translocation across membranes. These proteins interact with hydrophobic regions within target proteins via a carboxyl (C)-terminal peptide binding domain, with substrate binding and release being controlled 25 by the N-terminal ATP-binding domain of Hsc70/Hsp70. Hsc70/Hsp70-assisted folding reactions are accomplished by repeated cycles of peptide binding, refolding, and release,

which are coupled to ATP hydrolysis by the ATP-binding domain (ATPase) of Hsc70/Hsp70 and by subsequent nucleotide exchange. The chaperone activity of mammalian Hsc70/Hsp70 is regulated by partner proteins that either modulate the peptide binding cycle or that target the actions of these chaperones to specific proteins and subcellular compartments. DnaJ-family proteins (Hdj-1/Hsp40; Hdj-2; Hdj-3) stimulate the ATPase activity of Hsc70/Hsp70, resulting in the ADP-bound state which binds tightly to peptide substrates. The Hip protein collaborates with Hsc70/Hsp70 and DnaJ homologues in stimulating ATP hydrolysis, and thus also stabilize Hsc70/Hsp70 complexes with substrate polypeptides, whereas the Hop protein may provide co-chaperone functions through interactions with the C-terminal peptide binding domain.

The Bcl-2 associated athanogene-1 (bag-1) is named from the Greek word *athanos*, which refers to anti-cell death. BAG-1 was previously referred to as Bcl-2-associated protein-1 (BAP-1) in U.S. Patent No. 5,539,094 issued July 23, 1996, which is incorporated herein by reference. In this earlier patent, BAG-1 is described as a portion of the human BAG-1 protein, absent the N-terminal amino acids 1 to 85. In addition, a human protein essentially identical to human BAG-1 was described by Zeiner and Gehring, (*Proc. Natl. Acad. Sci., USA* **92**:11465-11469 (1995)). Subsequent to the issuance of U.S. Patent 5,539,094 the N-terminal amino acid sequence from 1 to 85 of human BAG-1 was reported.

BAG-1 and its longer isoforms BAG-1M (Rap46) and BAG-1L are recently described Hsc70/Hsp70-regulating proteins. BAG-1 competes with Hip for binding to the Hsc70/Hsp70 ATPase domain and promotes substrate release. BAG-1 also reportedly stimulates Hsc70-mediated ATP

hydrolysis by accelerating ADP/ATP exchange, analogous to the prokaryotic GrpE nucleotide exchange protein of the bacterial Hsc70 homologue, DnaK. Gene transfection studies indicate that BAG-1 proteins can influence a wide variety 5 of cellular phenotypes through their interactions with Hsc70/Hsp70, including increasing resistance to apoptosis, promoting cell proliferation, enhancing tumor cell migration and metastasis, and altering transcriptional activity of steroid hormone receptors.

10 Despite the notable progress in the art, there remains an unmet need for the further identification and isolation of additional homologous BAG protein species, and the nucleic acid molecules and/or nucleotide sequences that encode them. Such species would provide additional 15 means by which the identity and composition of the BAG domain, that is, the portion of the protein that is influencing or modulating protein folding, could be identified. In addition, such species would be useful for identifying agents that modulate apoptosis as candidates 20 for therapeutic agents, in particular, anticancer agents. The present invention satisfies these need, as well as providing substantial related advantages.

SUMMARY OF THE INVENTION

25 The present invention provides a family of BAG-1 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO: 4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)] , the invertebrate *C.elegans* [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast *S.pombe* 30 [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)] and the nucleic acid molecules that encode them.

Another aspect of the present invention provides an amino acid sequence present in the family of BAG-1 related proteins, that modulates Hsc70/Hsp70 chaperone activity, that is, the BAG domain.

5 Another aspect of the present invention provides novel polypeptide and nucleic acid compositions and methods useful in modulating Hsc70/Hsp70 chaperone activity.

Another aspect of the present invention is directed to methods for detecting agents that modulate the 10 binding of the BAG family of proteins, such as BAG-1 (beginning at residue 116 of SEQ ID NO:2), and related proteins with the Hsc70/Hsp70 Family of proteins or with other proteins that may interact with the BAG-Family proteins.

15 Still another aspect of the present invention is directed to methods for detecting agents that induce the dissociation of a bound complex formed by the association of BAG-Family proteins with Hsc70/Hsp70 Family molecule chaperones or other proteins.

20

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the full length cDNA sequence for human BAG-1 (SEQ ID NO:1) protein with the corresponding amino acid sequence (SEQ ID NO:2). Within the full length sequence are included the overlapping sub-sequences of 25 BAG-1 (beginning at nucleotide 391), BAG-1M [beginning at nucleotide 260 of (SEQ ID NO:2)], and BAG-1L [beginning at nucleotide 46 of (SEQ ID NO:2)].

Figures 2A and 2B combined shows the full length cDNA sequence (SEQ ID NO:3) aligned with the corresponding amino acid residues for human BAG-2 protein (SEQ ID NO:4).

Figure 3 shows a cDNA sequence (SEQ ID NO:5) aligned with the corresponding amino acid residues for human BAG-3 protein (SEQ ID NO:6).

Figure 4 shows the a cDNA sequence (SEQ ID NO:7) aligned with the corresponding amino acid residues for human BAG-4 protein (SEQ ID NO:8).

10 Figure 5 shows a cDNA sequence (SEQ ID NO:9) aligned with the corresponding amino acid residues for human BAG-5 protein (SEQ ID NO:10).

Figure 6A shows the full length cDNA sequence for *C. elegans* BAG-1 protein (SEQ ID NO:11).

15 Figure 6B shows the 210 amino acid sequence for *C. elegans* BAG-1 protein (SEQ ID NO:12).

Figure 7A shows the full length cDNA sequence for *C. elegans* BAG-2 protein (SEQ ID NO:13).

20 Figure 7B shows the 458 amino acid sequence for *C. elegans* BAG-2 protein (SEQ ID NO:14).

Figure 8A shows the full length cDNA sequence for *S. pombe* BAG-1A protein (SEQ ID NO:15).

Figure 8B shows the 195 amino acid sequence for *S. pombe* BAG-1A protein (SEQ ID NO:16).

Figure 9A shows the full length cDNA sequence for *S. pombe* BAG-1B protein (SEQ ID NO:17).

Figure 9B shows the 206 amino acid sequence for *S. pombe* BAG-1B protein (SEQ ID NO:18).

5 Figure 10 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10); *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). (A) The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like nuclear localization sequence are also shown. (B) The amino acid sequences of the BAG domain for human BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), BAG-4 (SEQ ID NO:8), BAG-5 (SEQ ID NO:10), *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18), and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) are aligned demonstrating 10 their homology. Black and gray shading represent identical 15 and similar amino acids, respectively.

Figure 11 shows assays demonstrating the interaction of BAG-family proteins with Hsc70/ATPase. (A) Two-hybrid assays using yeast expressing the indicated 25 fusion proteins. Blue color indicates a positive interaction, resulting in activation of the lacZ reporter gene. (B) *In vitro* protein assays using GST-fusion proteins and ³⁵S-labeled *in vitro* translated proteins. (C) Co-immunoprecipitation assays using anti-Flag or IgG1 30 control antibodies and lysates from 293T cells expressing Flag-tagged BAG-1 (beginning at residue 116 of SEQ ID

NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6), Daxx, or Apaf-1.

Figure 12 shows surface plasmon resonance analysis of BAG-family protein interactions with 5 Hsc70/ATPase. (A) SDS-PAGE analysis of purified recombinant proteins. (B) Representative SPR results of biosensor chips containing immobilized BAG proteins with and without maximally bound Hsc70/ATPase.

Figure 13 shows representative SPR results for 10 biosensor chips containing immobilized BAG-1 (beginning at residue 116 at SEQ ID NO:2), BAG-1(ΔC), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) proteins. Hsc70/ATPase was flowed over the chips (arrow/left) until maximal binding was reached (response units), then flow was continued without 15 Hsc70/ATPase (arrow/right). For BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6), Hsc70 was injected at 0.0175, 0.035, 0.07, 0.14, and 0.28 μM.

Figure 14 shows BAG-family protein modulation of 20 Hsc70 chaperone activity. (A) Protein refolding assay of chemically-denatured luciferase by Hsc70 plus DnaJ in the absence or presence of BAG and BAG-mutant proteins. (B) Concentration-dependent inhibition of Hsc70-mediated 25 protein refolding by BAG-family proteins [BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6)] but not by BAG-mutant (BAG-1 (ΔC)). (C) Hsc70/Hsp40-mediated refolding of heat-denatured luciferase was assayed in the presence of (black bars) or absence of (striped bars) of 1.8 μM Hip, with (lanes 3-10) or without 30 (lanes 1,2) various BAG-family proteins (1.8μM) as indicated (mean ±SE; n=3). A control (CNTL) is shown (lane 1) in which Hsc70 was replaced with an equivalent amount of BSA.

Figure 15A shows an expanded cDNA sequence for human BAG-3 protein (SEQ ID NO:19).

Figure 15B shows the corresponding amino acid residues for the human BAG-3 protein (SEQ ID NO:20) of 5 Figure 15A.

Figure 15C shows the expanded cDNA sequence (SEQ ID NO:19) aligned with the corresponding amino acid residues for human BAG-3 protein of Figure 15A (SEQ ID NO:20).

10 Figure 16A shows an expanded cDNA sequence for human BAG-4 protein (SEQ ID NO:21).

Figure 16B shows the corresponding amino acid residues for the human BAG-4 protein of Figure 16A (SEQ ID NO:22).

15 Figure 16C shows the expanded cDNA sequence (SEQ ID NO:21) aligned with the corresponding amino acid residues for human BAG-4 protein of Figure 16A (SEQ ID NO:22).

20 Figure 17A shows an expanded cDNA sequence for human BAG-5 protein (SEQ ID NO:23).

Figure 17B shows the corresponding amino acid residues for the human BAG-5 protein of Figure 17A (SEQ ID NO:24).

25 Figure 17C shows the expanded cDNA sequence (SEQ ID NO:23) aligned with the corresponding amino acid residues for human BAG-5 protein of Figure 17A (SEQ ID NO:24).

Figure 18 shows the topologies of the BAG-family proteins; human BAG proteins, BAG-1 (SEQ ID NO:2), BAG-2 (SEQ ID NO:4), expanded BAG-3 (SEQ ID NO:20), expanded BAG-4 (SEQ ID NO:22), expanded BAG-5 (SEQ ID NO:24); 5 *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and *C. elegans* BAG-1 (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14). The relative positions of the BAG domains are shown in black, ubiquitin-like regions are represented in gray, WW domain are represented in strips. Nucleoplasmin-like 10 nuclear localization sequence are also shown.

Definitions

The term "apoptosis", as used herein, refers to the process of programmed cell death, although not all programmed cell deaths occur through apoptosis, as used 15 herein, "apoptosis" and "programmed cell death" are used interchangeably.

The term "tumor cell proliferation", as used herein refers to the ability of tumor cells to grow and thus expand a tumor mass.

20 The term "cell migration", as used herein refers to the role cell motility plays in the invasion and potentially metastasis by tumor cells.

The term "metastasis", as used herein refers to the spread of a disease process from one part of the body 25 to another, as in the appearance of neoplasms in parts of the body remote from the site of the primary tumor; results in dissemination of tumor cells by the lymphatics or blood vessels or by direct extension through serous cavities or subarachnoid or other spaces.

The term "steroid hormone receptor function", as used herein refers to physiological, cellular and molecular functioning of receptors sites that bind with steroid hormones.

5 The term "substantially purified", as used herein, refers to nucleic acid or amino acid sequence that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with
10 which they are naturally associated.

"Nucleic acid molecule" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single or double stranded,
15 and represent the sense or antisense strand.

"Hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

20 The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, the sequence "A-G-T binds to the complementary sequence "T-C-A".

25 The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid and is referred to using the functional term
30 "substantially homologous." The inhibition of

hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A 5 substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency.

The term "antisense", as used herein, refers to 10 nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of 15 interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further 20 transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense, and "positive" is sometimes used in reference to the sense strand.

25 "Amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited herein this term excludes an amino 30 acid sequence of a naturally occurring protein. "Amino acid sequence", "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

The term "functional fragments" or "fragments", as used herein, with regard to a protein refers to portions of that protein that are capable of exhibiting or carrying out the activity exhibited by the protein as a whole. The 5 portions may range in size from three amino acid residues to the entire amino acid sequence minus one amino acid. For example, a protein "comprising at least a functional fragment of the amino acid sequence of SEQ ID NO:1", encompasses the full-length of the protein of SEQ ID NO:1 10 and portions thereof.

A "derivative" of a BAG protein, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The derivative may have "conservative" changes, wherein a substituted amino acid has similar 15 structural or chemical properties, e.g., substitution of an apolar amino acid with another apolar amino acid (such as replacement of leucine with isoleucine). The derivative may also have "nonconservative" changes, wherein a substituted amino acid has different but sufficiently 20 similar structural or chemical properties that permits such a substitution without adversely effecting the desired biological activity, e.g., replacement of an amino acid with an uncharged polar R group with an amino acid with an apolar R group (such as replacement of glycine with 25 tryptophan), or alternatively replacement of an amino acid with a charged R group with an amino acid with an uncharged Polar R group (such as replacement of lysine with asparagine).

Amino Acids - Apolar R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
alanine	methyl	ala	A
valine	2-propyl	aal	V
leucine	2-methylpropyl	leu	L
isoleucine	2-butyl	ile	I
proline	propyl* - cyclized	pro	P
phenylalanine	benzyl	phe	F
tryptophan	3-indolylmethl	tyr	W
methionine	methylthioethyl	met	M

Amino Acids - Uncharged Polar R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
glycine	H	gly	G
serine	hydroxymethyl	ser	S
threonine	1-hydroxyethyl	thr	T
cysteine	thiolmethyl	cys	C
tyrosine	4-hydroxyphenylmethyl	tyr	Y
asparagine	aminocarbonylmethyl	asn	N
glutamine	aminocarbonylethyl	gln	Q

20 Amino Acids - Charged R Groups

Amino Acid	Radical	Abbreviations	
		3-Letter	1-Letter
aspartic acid	carboxymethyl	asp	D
glutamic acid	carboxyethyl	glu	E
lysine	4-aminobutyl	lys	K
arginine	3-guanylpropyl	arg	R
histidine	4-imidazoylmethyl	his	H

Similar minor modifications may also include amino acids deletions or insertions or both. Guidance in determining which amino acid residues may be modified as indicated above without abolishing the desired biological 5 functionality may be determined using computer programs well known in the art, for example, DNASTAR software. In addition, the derivative may also result from chemical modifications to the encoded polypeptide, including but not limited to the following, replacement of hydrogen by an 10 alkyl, acyl, or amino group; esterification of a carboxyl group with a suitable alkyl or aryl moiety; alkylation of a hydroxyl group to form an ether derivative. Further a derivative may also result from the substitution of a L- configuration amino acid with its corresponding D- 15 configuration counterpart.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of a protein/polypeptide or portions thereof (such as BAG-1) and, as such, is able to 20 effect some or all of the actions of BAG-1 protein.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene 25 agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al., *Anticancer Drug Des.* 8:53-63 (1993)).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a family of BAG-1 30 related proteins from humans [BAG-1L (SEQ ID NO:2), BAG-1S beginning at residue 116 of SEQ ID NO:2, BAG-2 (SEQ ID

NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO: 8; and (SEQ ID NO:22) and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24)], the invertebrate *C.elegans* [BAG-1 (SEQ ID NO:12), BAG-2 (SEQ ID NO:14)] and the fission yeast *S.pombe* [BAG-1A (SEQ ID NO:16), BAG-1B (SEQ ID NO:18)], specifically the full length amino acid sequences comprising human BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), and BAG-2 (SEQ ID NO:4) *C.elegans* BAG-1 (SEQ ID NO:12), and BAG-2 (SEQ ID NO:14), and *S.pombe* BAG-1A (SEQ ID NO:16) and BAG-1B (SEQ ID NO:18); and partial sequences comprising human BAG-3 (SEQ ID NO: 6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) and functional fragments thereof. In particular, the invention provides the amino acid sequences comprising human BAG-2 (SEQ ID NO:4), BAG-3 (SEQ ID NO:6) and (SEQ ID NO:20), BAG-4 (SEQ ID NO:8) and (SEQ ID NO:22), and BAG-5 (SEQ ID NO:10) and (SEQ ID NO:24) proteins.

Another aspect of the present invention provides the nucleic molecule and nucleotide sequences that encode the family of BAG-1 related proteins from humans [BAG-1 (SEQ ID NO:1), BAG-2 (SEQ ID NO:3), BAG-3 (SEQ ID NO:5) and (SEQ ID NO:19), BAG-4 (SEQ ID NO:7) and (SEQ ID NO:21) and BAG-5 (SEQ ID NO:9) and (SEQ ID NO:23)], the invertebrate *C.elegans* [BAG-1 (SEQ ID NO:11), BAG-2 (SEQ ID NO:13)] and the fission yeast *S.pombe* [BAG-1A (SEQ ID NO:15), BAG-1B (SEQ ID NO:17)].

BAG-1L (SEQ ID NO:2) is a multifunctional protein that blocks apoptosis, promotes tumor cell metastasis, and contributes to factor-independent and p53-resistant cell growth. BAG-1L (SEQ ID NO:2) interacts with several types of proteins, including Bcl-2, some tyrosine kinase growth

factor receptors, steroid hormone receptors, and the p53-induced cell cycle regulator Siah-1A.

BAG-1 is a regulator of Hsc70/Hsp70 family molecular chaperones. A carboxyl-terminal domain in this 5 protein binds tightly to the ATPase domains of Hsc70 and Hsp70 ($K_d = 1$ nM) (Zeiner, M., Gebauer, M., and Gehring, U., *EMBO J.* **16**: 5483-5490, (1997)). BAG-1 modulates the activity of these molecular chaperones, acting as an apparent functional antagonist of the Hsp70/Hsc70-associated protein Hip (3-5) (Höhfeld, J. and Jentsch, S., *EMBO J.* **16**: 6209-6216, (1997); Takayama, S., Bimston, D. N., Matsuzawa, S., Freeman, B. C., Aime-Sempe, C., Xie, Z., Morimoto, R. J., and Reed, J. C., *EMBO J.* **16**: 4887-96, (1997); Zeiner, M., Gebauer, M., and Gehring, U., *EMBO J.* **16**: 5483-5490, (1997)). In general, protein refolding is 10 accomplished by Hsp70/Hsc70 through repeated cycles of target peptide binding and release, coupled to ATP hydrolysis (Ellis, R., *Curr Biol.* **7**: R531-R533, (1997)). BAG-1 appears to promote substrate release, whereas Hip 15 stabilizes Hsp70/Hsc70 complex formation with target peptides (Höhfeld, J., Minami, Y., and Hartl, F.-U., *Cell.* **83**: 589-598, (1995)). Since each substrate interaction with Hsc70/Hsp70 is unique in terms of the optimal length 20 of time the protein target should remain complexed with Hsc70/Hsp70 for achieving new conformations, the net effect 25 of BAG-1 can be either enhancement or inhibition of the refolding reaction.

The 70kd heat shock family proteins (Hsp70/Hsc70) are essential to a variety of cellular processes and have 30 been implicated in cancer, yet it is unclear how these proteins are regulated *in vivo*. A variety of co-chaperones have been identified which may target Hsp70/Hsc70 to different subcellular compartments or promote their

interactions with specific protein or protein complexes. BAG-1 appears to represent a novel Hsp70/Hsc70 regulator which differs functionally from all other mammalian co-chaperones identified to date, such as members of the 5 DnaJ-, Hip-, Hop-, and cyclophilin-families of proteins.

Another aspect of the present invention provides the amino acid sequence of a binding domain of about 40 to 55 amino acids that bind the a Hsc70/Hsp70 ATPase domain. The BAG domain is situated near the C-terminus, and the 10 ubiquitin-like domains are situated near the N-terminus.

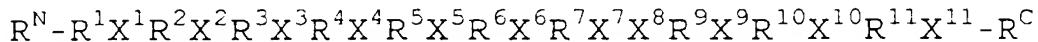
The BAG family of proteins of the present invention contain a common conserved C-terminal domain (the "BAG" domain) that facilitates binding to the ATPase domain of Hsp70/Hsc70. The carboxyl-terminal domain of BAG-1 15 binds to the ATPase domain of Hsc70/Hsp70 and regulates its chaperone function by acting as a ADP-ATP exchange factor. Other domains of BAG-1 mediate interactions with proteins such as Bcl-2 and retinoic acid receptors (RARs), allowing BAG-1 to target Hsc70/Hsp70 to other proteins, presumably 20 modulating their function by changing their conformations.

Human BAG-1 was previously shown to inhibit Hsc70/Hsp70 dependent refolding of denatured protein substrates *in vitro* (S. Takayama, et al., *EMBO J* 16, 4887-96 (1997); M. Zeiner, M. Gebauer, U. Gehring, *EMBO J.* 16, 25 5483-5490 (1997); and J. Höhfeld, S. Jentsch, *EMBO J.* 16, 6209-6216 (1997)). In Example III, Part A the effects of recombinant human BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were compared using *in vitro* protein refolding assays similar to those employed previously for assessing 30 BAG-1. The study showed that addition of equimolar amounts of each of the recombinant proteins to Hsc70 resulted in significant inhibition of luciferase refolding, with BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) showing somewhat

greater inhibitor activity than BAG-1 (Figure 4A). In a separate luciferase folding study BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) once again displayed inhibition of luciferase refolding, however in this study 5 varying amounts of BAG-1, BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) were added relative to Hsc70 which resulting in concentration-dependent inhibition of Hsc70 chaperone activity, i.e., luciferase folding (Example III Part A). Additional follow on studies using the same experimental 10 protocols as the previous studies, as taught in Example IIA, have shown that BAG-4 (SEQ ID NO:22) also undergoes association with Hsc70/ATPase.

Yet another aspect of the present invention provides a nucleotide sequence having at least about 15 15 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides that can hybridize or is complementary under relatively stringent conditions to a portion of the nucleic acid 20 sequences shown in Figures 1-9 and Figures 15-17, in particular the BAG domain as shown in Figure 1B, e.g., nucleotides 552-593 of human BAG-3, or nucleotides 167-221 of human BAG-4.

Yet another aspect of the present invention 25 provides a compound of the formula,



wherein,

R^N is a group of 1 to 552 independently selected amino acids;

30 R^1 is a group of 3 independently selected amino acids;

X^1 is an amino acid with a charged or uncharged R group, such as aspartic acid, glutamic acid, asparagine, or glutamine;

5 R^2 is a group of 7 independently selected amino acids;

X^2 is an amino acid with a charged R group, such as glutamic acid;

R^3 is a group of 5 independently selected amino acids;

10 X^3 is an amino acid with an apolar R group, such as leucine, methionine, or isoleucine;

R^4 is a group of 3 independently selected amino acids;

15 X^4 is an amino acid with charged R group, such as aspartic acid or glutamine acid;

R^5 is a single independently selected amino acid;

X^5 is an amino acid with apolar or uncharged R group, such as leucine, valine, methionine, alanine or threonine;

20 R^6 is a group of 15 independently selected amino acids;

X^6 is an amino acid with a charged or uncharged R group, such as arginine, lysine, glutamine or aspartic acid;

25 R^7 is a group of 2 independently selected amino acids;

X^7 is an amino acid with a charged R group, such as arginine;

30 X^8 is an amino acid with a charged R group, such as arginine or lysine;

R^9 is a group of 2 independently selected amino acids;

X^9 is an amino acid with an apolar R group, such as valine;

35 R^{10} is a group of 3 independently selected amino acids;

X^{10} is an amino acid with an uncharged R group, such as glutamine;

R^{11} is a group of 2 independently selected amino acids;

5 X^{11} is an amino acid with an apolar R group, such as leucine; and

R^C is a group of 1 to 100 independently selected amino acids.

A nucleotide sequence of at least about 15
10 nucleotides and, generally, about 25 nucleotides, preferably about 35 nucleotides, more preferably about 45 nucleotides, and most preferably about 55 nucleotides can be useful, for example, as a primer for the polymerase chain reaction (PCR) or other similar reaction mediated by
15 a polymerase such as a DNA or RNA polymerase (see PCR Protocols: A guide to methods and applications, ed. Innis et al. (Academic Press, Inc., 1990), which is incorporated herein by reference; see, for example, pages 40-41). In addition, such a nucleotide sequence of the invention can
20 be useful as a probe in a hybridization reaction such as Southern or northern blot analysis or in a binding assay such as a gel shift assay.

A nucleotide sequence of the invention can be particularly useful as an antisense molecule, which can be
25 DNA or RNA and can be targeted to all or a portion of the 5'-untranslated region or of the 5'-translated region of a bag-1 nucleic acid sequence in a cell. For example, an antisense molecule can be directed to at least a portion of the sequence shown as the BAG domain in Figure 1A, e.g.,
30 nucleotides 272-319 of human BAG-1L (SEQ ID NO:1), or nucleotides 79-147 of human BAG-5 (SEQ ID NO:9). Since the 5'-region of a nucleic acid contains elements involved in the control of expression of an encoded protein, an antisense molecule directed to the 5'-region of a nucleic

acid molecule can affect the levels of protein expressed in a cell.

A nucleotide sequence of the invention also can be useful as a probe to identify a genetic defect due a 5 mutation of a gene encoding a BAG protein in a cell. Such a genetic defect can lead to aberrant expression of a BAG protein in the cell or to expression of an aberrant BAG protein, which does not properly associate with a Bcl-2-related protein or Hsc70/Hsp70 protein in the cell. As a 10 result, a genetic defect in a gene encoding, for example, human BAG-1 can result in a pathology characterized by increased or decreased levels in protein folding.

Further a nucleotide compound or composition as taught in the present invention can be synthesized using 15 routine methods or can be purchased from a commercial source. In addition, a population of such nucleotide sequences can be prepared by restriction endonuclease or mild DNase digestion of a nucleic acid molecule that contains nucleotides as shown in the nucleotide sequences 20 shown in Figures 1-9 and Figures 15-17 that encodes the amino acids sequences also shown in Figures 1-9 and Figures 15-17. Methods for preparing and using such nucleotide sequences, for example, as hybridization probes to screen a library for homologous nucleic acid molecules 25 are well known in the art (see, for example, Sambrook et al., *Molecular Cloning: A laboratory manual* (Cold Spring Harbor Laboratory Press 1989); Ausubel et al., *Current Protocols in Molecular Biology* (Green Publ., NY 1989), each of which is incorporated herein by reference).

30 A particular nucleotide sequence can be designed based, for example, on a comparison of the nucleic acid molecules encoding any one of the BAG family proteins, as shown in Figures 1-9 and Figures 15-17, with another in the family. Such a comparison allows, for example, the

preparation of a nucleotide sequence that will hybridize to a conserved region present in both nucleic acid molecules, thus providing a means to identify homologous nucleic acid molecules present in other cell types or other organisms.

5 In addition, such a comparison allows the preparation of a nucleotide sequence that will hybridize to a unique region of any of the BAG family nucleotide sequences, such as those corresponding to the BAG domain, thus allowing identification of other proteins sharing this motif. In

10 this regard, it is recognized that, while the human BAG-3 proteins shown as Figures 3 and 20, and human BAG-5 proteins shown as Figures 5 and 24, are only partial sequences, a variant human BAG-3 or BAG-5 produced, for example, by alternative splicing can exist and can be

15 identified using an appropriately designed nucleotide sequence of the invention as a probe. Such useful probes readily can be identified by inspection of the sequences shown in the disclosed Figures by a comparison of the encoding nucleotide sequences.

20 If desired, a nucleotide sequence of the invention can incorporate a detectable moiety such as a radiolabel, a fluorochrome, a ferromagnetic substance, a luminescent tag or a detectable binding agent such as biotin. These and other detectable moieties and methods of

25 incorporating such moieties into a nucleotide sequence are well known in the art and are commercially available. A population of labelled nucleotide sequences can be prepared, for example, by nick translation of a nucleic acid molecule of the invention (Sambrook et al., *supra*,

30 1989; Ausubel et al., *supra*, 1989).

One skilled in the art would know that a method involving hybridization of a nucleotide sequence of the invention can require that hybridization be performed under relatively stringent conditions such that nonspecific

35 background hybridization is minimized. Such hybridization

conditions can be determined empirically or can be estimated based, for example, on the relative GC content of a sequence and the number of mismatches, if known, between the probe and the target sequence (see, for example, 5 Sambrook et al., *supra*, 1989).

The invention further provides antibodies specific for human BAG family protein. As used herein, the term "antibody" includes polyclonal and monoclonal antibodies, as well as polypeptide fragments of antibodies 10 that retain a specific binding activity for human BAG-1 of at least about $1 \times 10^5 \text{ M}^{-1}$. One skilled in the art would know that anti-BAG-1 antibody fragments such as Fab, F(ab')₂ and Fv fragments can retain specific binding activity for human BAG-1 (beginning at residue 116 of SEQ ID NO:2) and, 15 thus, are included within the definition of an antibody. In addition, the term "antibody" as used herein includes naturally occurring antibodies as well as non-naturally occurring antibodies and fragments that retain binding activity such as chimeric antibodies or humanized 20 antibodies. Such non-naturally occurring antibodies can be constructed using solid phase peptide synthesis, can be produced recombinantly or can be obtained, for example, by screening combinatorial libraries consisting of variable heavy chains and variable light chains as described by Huse 25 et al., *Science* 246:1275-1281 (1989), which is incorporated herein by reference.

One skilled in the art would know that purified BAG family protein, which can be prepared from natural sources or synthesized chemically or produced 30 recombinantly, or portions of a BAG family protein, including a portion of human BAG family protein such as a synthetic peptide as described above, can be used as an immunogen. Such peptides useful for raising an antibody include, for example, peptide portions of the N-terminal 85 35 amino acids or the BAG domain of any of the human BAG

proteins (see Figure 1B). A particularly advantageous use of such a protein is for the immunostaining, wherein the methods provides a process to contrast the immunostaining of BAG-family proteins in carcinoma cells with adjacent 5 non-neoplastic prostatic epithelial and basal cells which are generally present in the same tissue sections. These results would be correlated with a Gleason grade to determine whether any of the BAG-family proteins tend to be expressed at higher or lower levels in histologically 10 advanced tumors. From this process a determination can be made as to degree at which the disease is progressing in a given patient, i.e., a prognosis can be made.

Non-immunogenic fragments or synthetic peptides of BAG proteins can be made immunogenic by coupling the 15 hapten to a carrier molecule such bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH), as described in Example IV, below. In addition, various other carrier molecules and methods for coupling a hapten to a carrier molecule are well known in the art and described, for 20 example, by Harlow and Lane, *Antibodies: A laboratory manual* (Cold Spring Harbor Laboratory Press, 1988), which is incorporated herein by reference.

EXAMPLES

The following examples are given to enable those 25 skilled in the art to more clearly understand and to practice the present invention. They should not be considered as limiting the scope of the invention, but merely as being illustrative and representative thereof.

EXAMPLE I

Isolation and Characterization
of BAG-family cDNA Sequences

This example describes methods for isolating and
5 characterizing of BAG-family cDNA sequences from human,
nematode and yeast.

A. Cloning of human BAG cDNA sequences

Yeast two-hybrid library screening of a human
Jurkat cell cDNA library was performed as described by
10 Takayama et al., EMBO J., 16:4887-96 (1997); Matsuzawa et
al., EMBO J., 17:2736-2747 (1998), which are incorporated
herein by reference) using EGY48 strain yeast transformed
with pGilda-Hsc70/ATPase (67-377 amino acids) and the lacZ
reporter plasmid pSH18-34. Of the resulting $\sim 5 \times 10^6$
15 transformants, 112 Leu^r colonies were obtained after
1 week incubation at 30°C. Assay of β -galactosidase (β -gal)
activity of these colonies resulted in 96 clones. Mating
tests were then performed using RFY206 yeast strain
transformed with pGilda, pGilda mBAG-1 (1-219), or pGilda
20 Hsc70/ATPase. Of these, 66 displayed specific interactions
with Hsc70/ATPase. The pJG4-5 cDNAs were recovered using
KC8 *E. coli* strain which is auxotrophic for tryptophan
(Trp). DNA sequencing revealed 3 partially overlapping
25 human BAG-1, 4 identical and one overlapping cDNAs encoding
BAG-2, and 2 partially overlapping BAG-3 clones.

Using the above described yeast two-hybrid screen
with the ATPase domain of Hsc70 as "bait", several human
cDNAs were cloned which encode portions of BAG-1 or of two
other BAG-1-like proteins which are termed BAG-2 (SEQ ID
30 NO:4) and BAG-3 (SEQ ID NO:6). The longest of the cDNAs
for BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) contained
open reading frames (ORFs) of 207 and 162 amino acids,
respectively, followed by stop codons. All BAG-1 (SEQ ID

NO:1), BAG-2 (SEQ ID NO:3) and BAG-3 (SEQ ID NO:5) cDNAs obtained by two-hybrid library screening with Hsc70/ATPase contained a conserved domain of about 40-50 amino acids which are termed the "BAG" domain and are shown in Figure 5 10. These results demonstrate that a family of BAG-1-related proteins all contain a conserved ~45 amino acid region near their C-terminus that binds Hsc70/Hsp70.

B. Identification of additional BAG-family proteins

A search of the translated Genbank database using 10 the bBLAST and FASTA search programs also identified human ESTs that provided sequences for further investigation of BAG-family proteins. The putative BAG-4 (SEQ ID NO:8) and BAG-5 (SEQ ID NO:10) proteins contain BAG-domains that share the greatest sequence similarity with the BAG-domain 15 of BAG-3 (SEQ ID NO:6). These were designated BAG-4 (Accession number AA693697, N74588) and BAG-5 (Accession number AA456862, N34101). BAG-4 has 62% identity and ~81% similarity to BAG-3, and BAG-5 has 51% identity and ~75% similarity to BAG-3.

20 Additional BAG-family orthologues or homologues were also identified using computer-based searches and resulted in BAG-family homologue in the nematode *C. elegans* and the fission yeast *S. pombe*. The *C. elegans* genome encodes two apparent BAG-family proteins, which are most 25 similar in their overall sequences to the human BAG-1 (Afo39713, gi:2773211) (SEQ ID NO:12) and BAG-2 (SEQ ID NO:14) (Afo68719, gi:3168927). The *S. pombe* contains two BAG-family proteins that share the greatest overall sequence similarity with human BAG-1 (Alo23S54, gi/3133105 30 and Alo23634, gi/3150250). The human and *C. elegans* BAG-1 proteins as well as *S. pombe* BAG-1A all have ubiquitin-like domains near their N-termini (see Figure 10A) of unknown function.

The overall predicted amino acid sequences of the *C. elegans* BAG-1 (SEQ ID NO:12) and *S. pombe* BAG-1A (SEQ ID NO:16) proteins are ~18% identical (~61% similar) and ~17% identical (~64% similar), respectively, to human BAG-1, 5 implying origin from a common ancestral gene. The *C. elegans* BAG-1 protein (SEQ ID NO:12), however, contains a 5 to 7 amino acid insert in its BAG-domain as compared to the human, murine, and yeast BAG-1 homologues (see Figure 10B), and is more similar to BAG-2 (SEQ ID NO:4) in regard 10 to its BAG-domain. *C. elegans* and human BAG-2 also may be derived from a common ancestor as the C-terminal 225 amino acid region which encompasses both the BAG domain and upstream region of both *C. elegans* and human BAG-2 share ~34% amino acid sequence identity and ~70% similarity. The 15 human BAG-2 protein (SEQ ID NO:4), however, contains a 9 amino acid insert in its BAG-domain compared to its *C. elegans* counterpart (see Figure 10B). Evolutionary-tree prediction algorithms suggest that human and *C. elegans* BAG-2 represent a distinct branch of the BAG-family that is 20 more evolutionarily distant from the other BAG-family proteins. None of the predicted BAG-family proteins contain recognizable regions analogous to those found in other Hsc70 regulatory proteins, such as the J-domains and G/F-domains of DnaJ family proteins and the 25 Tetratricopeptide Repeat (TR) domains of Hip/Hop family proteins.

C. Yeast two-hybrid assay of BAG binding to Hsc70/ATPase

The longest of the cDNAs obtained for the BAG-2 and BAG-3 proteins were expressed with N-terminal 30 transactivation (TA) domains in yeast and tested by yeast two-hybrid assay for interactions with fusion proteins consisting of Hsp70/ATPase or a variety of unrelated proteins (Fas, Siah, Fadd) containing N-terminal LexA DNA-binding domains. TA-BAG-2 and TA-BAG-3 demonstrated

positive interactions with LexA-Hsc70/ATPase, resulting in transactivation of a *lacZ* reporter gene that was under the control of LexA operators (Figure 11A). No interactions with LexA-Fas (cytosolic domain), LexA-Siah, LexA-Fadd, or 5 LexA were detected (see Figure 11A) demonstrating that the BAG-2 and BAG-3 proteins interact specifically with Hsc70/ATPase. Specific two-hybrid interactions between Hsc70/ATPase and either BAG-2 or BAG-3 were also observed when BAG-2 and BAG-3 were expressed as LexA DNA-binding 10 domain fusion proteins and Hsc70/ATPase was fused with a TA domain (see Figure 11A; right panel). These results demonstrate that similarly to BAG-1, BAG-2 and BAG-3 specifically interact with Hsc70/ATPase.

15 In order to determine whether the BAG proteins are capable of forming heterodimers, coexpression of BAG-2 and BAG-3 in the yeast two-hybrid assay was also performed. Coexpression of BAG-2 and BAG-3 failed to show interaction with BAG-1 or a deletion mutant of BAG-1 (Δ C) which is missing part of its C-terminal domain required for 20 Hsp70/Hsc70 binding suggest that these proteins do not form heterodimers.

D. Isolation and characterization of the complete open reading frame sequences of BAG-2 and BAG-3

25 In order to deduce the complete ORFs of BAG-2 and BAG-3, a λ -phage cDNA library was screened as follows, using hybridization probes derived from the two-hybrid screening. A human jurkat T-cell λ -ZapII library cDNA library (Stratagene) was screened by hybridization using 32 P-labeled purified insert DNA from the longest of the 30 human BAG-2 (clone #11) and human BAG-3 (clone #28) cDNA clones. From about one million clones screened, 38 BAG-2 and 23 BAG-3 clones were identified, cloned, and their cDNA inserts recovered as pSKII plasmids using a helper phage method (Stratagene). DNA sequencing of λ -phage derived

human BAG-2 cDNA clones revealed an ORF encoding a predicted 211 amino acid protein, preceded by an in-frame stop codon. The longest human BAG-3 λ -phage cDNA clone contains a continuous ORF of 682 amino acids followed by a 5 stop codon, but without an identifiable start codon (see Figure 10A).

Although BAG-1L (SEQ ID NO:2), BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) all contain a homologous BAG domain 10 near their C-terminus, the N-terminal regions of these proteins are dissimilar. Using a combination of search tools (Prosite Search: PP search, using the Prosite pattern database, BCM Search Launcher, Baylor College of Medicine, and Blocks Search), it was determined that the BAG-2 N- 15 terminal region contains potential kinase phosphorylation sites but otherwise shares no apparent similarity with other proteins or known functional domains.

In contrast, the predicted N-terminal region BAG-3 contains a WW domain as shown in Figure 10A. WW 20 domains have been identified in a wide variety of signaling proteins, including a Yes kinase adaptor protein (YAP), the Na⁺-channel regulator Nedd4, formin-binding proteins, dystrophin, and the peptidyl prolyl cis-trans-isomerase Pin-1. These roughly 40 amino acid domains mediate protein 25 interactions and bind the preferred peptide ligand sequence xPPxY (Sudol., TIBS, 21: 161-163, 1996, which is incorporated herein by reference).

EXAMPLE IIIn vitro Association of
BAG proteins and Hsc70/ATPase

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) bind Hsc70/ATPase in various *in vitro* assays.

A. Solution binding assay of BAG-2 and BAG-3 to Hsc70/ATPase

Association of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) with Hsc70/ATPase was determined by an *in vitro* protein binding assay where Hsc70/ATPase or BAG-family proteins were expressed in bacteria as Glutathione S-Transferase (GST) fusion proteins. Purified cDNA sequences encoding residues 5 to 211 of human BAG-2 (clone #11) and the C-terminal 135 amino acids of human BAG-3 (clone #28) (see Figure 10A) were subcloned into the EcoRI/Xho I sites of pGEX4T-1 prokaryotic expression plasmid (Pharmacia; Piscataway, NJ). These plasmids as well as pGEX4T-1-BAG-1, pGEX-4T-1-BAG-1 (Δ C), and pGEX-4T-1-XL which have been described previously (Takayama et al., *supra* (1997); Xie et al., *Biochemistry*, 37:6410-6418, (1998), which are incorporated herein by reference), were expressed in XL-1 blue strain *E. Coli* (Stratagene, Inc., La Jolla, CA). Briefly, a single colony was inoculated into 1L of LB media containing 50 μ g/ml ampicillin and grown at 37°C overnight. The culture was then diluted by half with fresh LB/ampicillin and cooled to room temperature for 1 hr, before inducing with 0.4mM IPTG for 6 h at 25°C.

Cells were recovered and incubated with 0.5 mg/ml lysozyme in 50 mM Tris (pH 8.0), 150 mM NaCl, 1% Tween-20, 0.1% 2-mercaptoethanol, 5 mM EDTA, 1 mM PMSF and a mixture

of other protease inhibitors obtained from Boehringer Mannheim (1697498) at room temperature for 0.5 h, followed by sonication. Cellular debris were pelleted by centrifugation at 27,500g for 10 min and the resulting supernatants were incubated with 30 ml of glutathione-Sepharose (Pharmacia) at 4°C overnight. The resin was then washed with 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, and 0.1% 2-mercaptoethanol until the OD 280nm reached <0.01. For removal of GST, the resin with immobilized GST-fusion protein was incubated with 10U of thrombin (Boehringer, Inc.) at 4°C in 20 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Tween-20, 0.1% 2-Mercaptoethanol, and 2.5 mM CaCl₂ overnight. Released proteins were then purified on Mono Q (HR10/10, Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0 and dialyzed into chaperone assay buffer.

The ability of BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) to bind Hsc70/ATPase in solution was then examined. GST control or GST-BAG proteins were immobilized on glutathione-Sepharose and tested for binding to ³⁵S-labeled *in vitro* translated (IVT) proteins. Immunoprecipitation and *in vitro* GST-protein binding assays were performed as described by Takayama et al., *supra* (1997), using pCI-Neo flag or pcDNA3-HA into which human Bag-2 (clone #11) or human BAG-3 (clone #28) had been subcloned for *in vitro* translation of ³⁵S-L-methionine labeled proteins or expression in 293T cells. As shown in Figure 11B, ³⁵S-Hsc70/ATPase bound *in vitro* to GST-BAG-1, GST-BAG-2, and GST-BAG-3 but not to GST-BAG-1(ΔC) or several other control proteins. BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) also exhibited little or no binding to themselves or to each other, demonstrating that these proteins do not strongly homo- or hetero-dimerize or oligomerize. It should be noted, however, that BAG-2 (SEQ

ID NO:4) displayed weak interactions with itself in binding assays and produced a positive result in yeast two-hybrid experiments, demonstrating that it can have the ability to self-associate.

5 B. Binding of BAG proteins to Hsc70 in vivo

The ability of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins to interact in cells with Hsc70 was tested by expressing these proteins with N-terminal Flag epitope tags in 293T human epithelial cells using co-10 immunoprecipitation assays as described previously (Takayama et al., *supra* (1997)). cDNAs encoding the λ -phage cloned regions of BAG-2 and BAG-3 were subcloned in-frame into pcDNA3-Flag. Anti-Flag immune complexes prepared from 293T cells after transfection with plasmids 15 encoding Flag-BAG-1, Flag-BAG-2, or Flag-BAG-3 were analyzed by SDS-PAGE/immunoblot assay. As shown in Figure 10C, antiserum specific to Hsc70 detected the presence of BAG proteins associated with Hsc70, whereas control immune-complexes prepared with IgG1 as well as anti-Flag immune 20 complexes prepared from cells transfected with Flag-tagged control proteins, Daxx and Apaf-1, did not contain Hsc70 associated protein. These results further demonstrate that BAG-family proteins specifically bind to Hsc70.

25 C. BIACore assay of BAG protein binding to the ATPase domain of Hsc70

BAG-1 (beginning at residue 116 of SEQ ID NO:2) is known to bind tightly to the ATPase domain of Hsc70 (Stuart et al., *J. Biol. Chem.*, In Press (1998)). BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins were 30 therefore, examined for their ability to bind to Hsc70/ATPase. The affinity and binding kinetics of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) to Hsc70/ATPase was also compared to that of BAG-1 (beginning at residue 116 of

SEQ ID NO:2) for Hsc70/ATPase, using a surface plasmon resonance technique (BIAcore) which has been described previously (Stuart et al., *supra*, (1998) which is incorporated herein by reference).

5 BAG-family proteins were produced in bacteria and purified to near homogeneity as shown in Figure 12A and described above in Example I. The purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), -2 (SEQ ID NO:4), and -3 (SEQ ID NO:6) proteins were then immobilized
10 on biosensor chips and tested for their interactions with Hsc70 in the soluble phase. Kinetic measurements were performed using a BIAcore-II instrument with CM5 sensor chip and Amine Coupling Kit (Pharmacia Biosensor AB, Sweden). Briefly, for immobilization of proteins, the
15 sensor chip was equilibrated with HK buffer (10 mM Hepes (pH 7.4), 150 mM KCL) at 5 μ l/min, then activated by injecting 17 μ l of 0.2M N-ethyl-N'-(3-diethylaminopropyl)-carbodiimide and 0.05M N-hydroxysuccinimide (NHS/EDC) followed by 35 μ l of the protein of interest, in 10 mM acetate, pH 3.5-4.5. Excess NHS-ester on the surface was
20 deactivated with 17 μ l 1M ethanolamine-HCL (pH8.5). After immobilization, 5 μ l of regeneration buffer (50 mM phosphate (pH 6.8) and 4M GuHCl) was injected. For binding assays, Hsp70 (Sigma, H8778) was dissolved in HK buffer, and
25 injected at 10 μ l/min across the prepared surface at various concentrations. The surface was regenerated after each injection with 5 μ l of regeneration buffer. The rate constants k_{ass} and k_{diss} were generated with BIAevaluation softward 3.01 (Pharmacia Biosensor AB). Addition of Hsc70
30 to chips containing BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4) or BAG-3 (SEQ ID NO:6) resulted in concentration-dependent binding, as reflected by an increase in the Response Units (RU) measured at the chip surface (shown in Figure 3B). In contrast, Hsc70
35 failed to display interactions in BIAcore assays with a variety of control proteins as well as a mutant of BAG-1

lacking a C-terminal portion of the BAG domain which is required for Hsc70-binding (Figure 3B). Furthermore, flowing of various control proteins such as GST, BSA and Bcl-XL over the BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) chips resulted in negligible interaction. These results further demonstrate the specificity with which BAG-family proteins interact with and bind to Hsc70.

The rates of Hsc70 binding to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) proteins were similar, following pseudo first-order kinetics with estimated association rate constants (κ_a) of 2.1, 2.1 and $2.4 \times 10^5 \text{ M}^{-1} \text{ sec}^{-1}$, respectively. After allowing binding of Hsc70 to immobilized BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to reach plateau levels, the chaperone was removed from the flow solution and the dissociation rate was monitored. BAG-1 (beginning at residue 116 at SEQ ID NO:2) and BAG-2 (SEQ ID NO:4) exhibited similar dissociation rates, with relatively slow loss of Hsc70 from the chip surface, resulting in estimated dissociation rate constants (κ_d) of 3.0 and $5.0 \times 10^{-4} \text{ sec}^{-1}$, respectively (see Figure 3B). In contrast, Hsc70 dissociated more rapidly from biosensor chips containing BAG-3 (see Figure 3B), yielding an estimated κ_d of $1.7 \times 10^{-3} \text{ sec}^{-1}$. From the kinetic data, the apparent affinities (κ_D = κ_d/κ_a) were calculated for binding of Hsc70 to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), and BAG-3 (SEQ ID NO:6) and were estimated to equal about $K_D = 1.4 \text{nM}$, $K_D = 2.4 \text{nM}$, and $K_D = 7.4 \text{nM}$, respectively. These results demonstrate that the interactions of BAG-family proteins with Hsc70 occur with apparent affinities sufficient for physiological relevance.

EXAMPLE III

BAG-family proteins inhibit
Hsp70/Hsc70-dependent protein folding

This example demonstrates that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) proteins inhibit Hsp70/Hsc70-dependent refolding of denatured proteins similarly to a BAG-1 (beginning at residue 116 of SEQ ID NO:2) protein.

The effects of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) protein on Hsp70/Hsc70-dependent protein refolding was determined using *in vitro* protein refolding assays similar to those described previously by Takayama et al., *supra*, 1998; Terada et al., *J Cell Biol.*, 139:1089-1095, 1997, which are incorporated herein by reference. Briefly, luciferase (20 μ M) was denatured in 25 mM Hepes-KOH, pH 7.2, 15 50 mM potassium acetate, 5 mM DTT, 6M guanidine hydrochloride at ~25°C for 1 h. Denatured luciferase was diluted 1:40 into 25 mM Hepes-KOH, pH 7.2, 50 mM potassium acetate, 5 mM DTT. Hsc70 (1.8 μ M), DnaJ (StressGen, Inc.) (0.9 μ M), and various purified recombinant proteins as 20 indicated were added to refolding buffer (30 mM Hepes-KOH, pH 7.6, 120 mM potassium acetate, 3mM magnesium acetate, 2 mM DTT, 2.5 mM ATP) with 0.2 volume of diluted denatured luciferase to a final concentration of 0.1 μ M. Luciferase activity was measured after 1.5 hr incubation at 35°C.

25 The combination of Hsc70 and DnaJ resulted in ATP-dependent refolding of chemically denatured firefly luciferase, with function of over half the denatured enzyme restored in a 90 minute reaction, as monitored by a chemiluminescence assay. In contrast, neither Hsc70 nor 30 DnaJ alone were able to induce substantial refolding of denatured luciferase. Furthermore, little spontaneous

restoration of luciferase activity was observed with control proteins, BSA, GST or Bcl-XL (see Figure 4A).

Addition of recombinant purified BAG-1 (beginning 5 at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) to the above assays in amounts equimolar to Hsc70 (1.8 μ M) resulted in striking inhibition of luciferase refolding. BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) displayed somewhat greater inhibitory 10 activity than BAG-1 (beginning at residue 116 of SEQ ID NO:2) as shown in Figure 4A. In contrast, the BAG-1 (Δ C) protein, which fails to bind Hsc70 as well as several other control proteins, had no effect on luciferase refolding.

In an additional refolding assay, described 15 previously by Minami et al., J Biol. Chem. 271:19617-24, 1996), purified Hsc70 and human DnaJ homolog Hdj-1 (Hsp 40) were used with additional cofactors provided in reticulocyte lysates (5% v:v) to produce a system capable of refolding denatured luciferase. Briefly, additional 20 cofactors included, recombinant Luciferase (Promega: QuantiLum TM), that had been heat denatured at 42°C for 10 min, 1.8 μ M Hsc70 (Sigma; purified from bovine brain), 0.9 μ M Hsp40, and various recombinant purified proteins. Luciferase activity was measured (Promega luciferase assay 25 kit) using a luminometer (EG&G Berthold, MicroLumat luminometer, Model #LB96P). All results were normalized relative to non-denatured luciferase that had been subjected to the same conditions. Control reactions lacking ATP, Hsc70, or Hsp40 resulted in negligible 30 luciferase refolding.

Various amounts of purified BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6), relative to amounts of Hsc70 were used in the above-described protein refolding assay. Addition of 35 BAG-family proteins resulted in a concentration-dependent

inhibition of Hsc70 chaperone activity. Furthermore, the BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) inhibition of Hsc70 chaperone activity was demonstrated to be as potent as that observed for BAG-1 (beginning at residue 116 of SEQ 5 ID NO:2). In contrast, the BAG-1 (Δ C) mutant as well as other control proteins did not suppress Hsc70-mediated refolding of denatured luciferase. These results indicate that BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) can inhibit Hsc70/Hsp70 dependent protein refolding activity to 10 the same extent as BAG-1 (beginning at residue 116 of SEQ ID NO:2).

B. BAG competes with Hip for binding to Hsc70.

It is known that BAG-1 competes with Hip for binding to Hsc70, with these proteins exerting opposite 15 effects on Hsc70-mediated protein refolding (Hohfeld, J., and Jentsch, S., *Embo J.*, 16:6209-6216, 1997, which is incorporated herein by reference). In order to determine whether BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6) also compete with Hip for binding to Hsc70, refolding assays 20 were performed as described above in the presence of Hip protein.

Hip was purified as His₆-protein. The fusion protein was induced from pET28-Hip (V. Prapapanich et al., *Mol Cell Biol.*, 18:944-952, 1998, which is incorporated 25 herein by reference) with 0.1 mM IPTG at 25°C for 6h in BL21 cells. Cells from 1L of culture were resuspended into 50 ml of 50 mM Phosphate buffer (pH 6.8), 150 mM NaCl, and 1% (v/v) Tween-20 and then incubated with 0.5 mg/ml lysozyme at 25°C for 0.5h, followed by sonication. After 30 centrifugation at 27,500g for 10 min, the resulting supernatant was mixed with 15 ml nickel resin (Qiagen, Inc.) at 4°C for 3 h with 25 mM imidazol. The resin was then washed with 50 mM phosphate buffer (pH 6.8), 25 mM imidazol, 150 mM NaCl and 0.1% Tween-20 until the OD_{280nm}

reached a value of <0.01. His₆-Hip protein was eluted with 250 mM imidazol in washing buffer (Qiagene, Inc.) and purified on Mono Q (HR10/10 Pharmacia) by FPLC using a linear gradient of 0.5M NaCl at pH 8.0, followed by 5 dialysis in chaperone assay buffer.

In the refolding assay reactions, addition of purified Hip at equimolar concentrations relative to BAG-1 (beginning at residue 116 of SEQ ID NO:2), BAG-2 (SEQ ID NO:4), or BAG-3 (SEQ ID NO:6) (1.8 μ M) completely negated 10 the inhibitory effects of the BAG-family proteins on refolding of denatured luciferase (see Figure 4C). These results demonstrate that the suppression of Hsc70 chaperone activity by BAG-family proteins is reversible, and that Hip antagonizes the effects of not only BAG-1 (beginning at 15 residue 116 of SEQ ID NO:2), but also of BAG-2 (SEQ ID NO:4) and BAG-3 (SEQ ID NO:6).

In summary, these results demonstrate that BAG-family proteins all contain a conserved BAG domain near their C-terminus that binds Hsc70/Hsp70, and that human 20 BAG-family proteins can bind with high affinity to the ATPase domain of Hsc70 and inhibit its chaperone activity through a Hip-repressable mechanism.

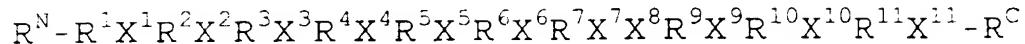
EXAMPLE IV

EXPANDED NUCLEIC ACID AND AMINO ACID SEQUENCES
25 FOR HUMAN BAG-3, BAG-4 AND BAG-5

Following the procedures disclosed herein, the nucleic acid and amino acids sequences to human BAG-3, BAG-4 and BAG-5 were further expanded. The expanded sequences for BAG-3, BAG-4 and BAG-5 are shown in 30 Figures 15, 16 and 17, respectively, with their respective sequence identification numbers, "SEQ ID NO"s.

We claim:

1. A compound of the formula,



wherein,

5 R^N is a group of about 1 to 552 independently selected amino acids;

10 R^1 is a group of 3 independently selected amino acids;

15 X^1 is an amino acid with a charged or uncharged R group;

20 R^2 is a group of 7 independently selected amino acids;

25 X^2 is an amino acid with a charged R group;

30 R^3 is a group of 5 independently selected amino acids;

35 X^3 is an amino acid with an apolar R group;

40 R^4 is a group of 3 independently selected amino acids;

45 X^4 is an amino acid with charged R group;

50 R^5 is a single independently selected amino acid;

55 X^5 is an amino acid with apolar or uncharged R group;

60 R^6 is a group of 15 independently selected amino acids;

65 X^6 is an amino acid with a charged or uncharged R group;

70 R^7 is a group of 2 independently selected amino acids;

75 X^7 is an amino acid with a charged R group;

80 X^8 is an amino acid with a charged R group;

85 R^9 is a group of 2 independently selected amino acids;

90 X^9 is an amino acid with an apolar R group;

R^{10} is a group of 3 independently selected amino acids;

X^{10} is an amino acid with an uncharged R group;

5 R^{11} is a group of 2 independently selected amino acids;

X^{11} is an amino acid with an apolar R group; and

R^C is a group of about 1 to 100 independently selected amino acids.

2. A substantially purified nucleic acid
10 molecule having a nucleotide sequence corresponding to or complementary to at least 20 nucleotides from a nucleotide sequence selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).

15 3. The nucleic acid of claim 2 having a nucleotide sequence corresponding to or complementary to a nucleotide sequence that encodes a functionally active BAG family protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).

20 4. The nucleic acid of claim 3 selected from the group consisting of (SEQ ID NO:1), (SEQ ID NO:3), (SEQ ID NO:5), (SEQ ID NO:7), (SEQ ID NO:9), (SEQ ID NO:19), (SEQ ID NO:21) and (SEQ ID NO:23).

25 5. The nucleic acid of claim 3 complementary to a nucleotide sequence that encodes a functionally active BAG protein selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24).

6. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:3).

7. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:5).

8. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:7).

5

9. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:9).

10. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:19).

10 11. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:21).

12. A substantially purified nucleic acid molecule having the nucleotide sequence of (SEQ ID NO:23).

13. A substantially purified BAG family protein
15 encoded by the nucleic acid molecule of claim 1.

14. A substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8), (SEQ ID NO:10), (SEQ ID NO:20), (SEQ
20 ID NO:22) and (SEQ ID NO:24) or a fragment, a derivative or a mimetic thereof.

15. A substantially purified protein corresponding to the amino acid sequence of 157 to 204 of (SEQ ID NO:2).

25 16. A substantially purified protein corresponding to the amino acid sequence of 272 to 319 of (SEQ ID NO:2).

17. A substantially purified protein corresponding to the amino acid sequence of 164 to 211 of (SEQ ID NO:4).

18. A substantially purified protein 5 corresponding to the amino acid sequence of 418 to 510 of (SEQ ID NO:20).

19. A substantially purified protein corresponding to the amino acid sequence of 378 to 457 of (SEQ ID NO:22).

10 20. A substantially purified protein corresponding to the amino acid sequence of 6 to 97 of (SEQ ID NO:24).

15 21. A substantially purified protein corresponding to the amino acid sequence of 180 to 257 of (SEQ ID NO:24).

22. A substantially purified protein corresponding to the amino acid sequence of 272 to 349 of (SEQ ID NO:24).

23. A substantially purified protein 20 corresponding to the amino acid sequence of 362 to 444 of (SEQ ID NO:24).

24. A pharmaceutical composition comprising a nucleic acid molecule of claim 1 useful for modulating tumor cell proliferation, cell migration and metastasis, 25 and steroid hormone receptor function.

25. A method of modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function by administering a nucleic acid molecule of claim 1.

26. A pharmaceutical composition comprising a substantially purified BAG family protein comprising of the amino acid sequence selected from the group consisting of (SEQ ID NO:2), (SEQ ID NO:4), (SEQ ID NO:6), (SEQ ID NO:8),
5 (SEQ ID NO:10), (SEQ ID NO:20), (SEQ ID NO:22) and (SEQ ID NO:24), or a fragment, a derivative or a mimetic thereof, useful for modulating tumor cell proliferation, cell migration and metastasis, and steroid hormone receptor function.

10 27. A method of modulating tumor cell proliferation by administering a pharmaceutical composition of claim 26.

15 28. A method of modulating cell migration and metastasis by administering a pharmaceutical composition of
claim 26.

29. A method of modulating steroid hormone receptor function by administering a pharmaceutical composition of claim 26.

30. A substantially purified antibody that
20 specifically binds to a BAG family protein of claim 14.

31. The antibody of claim 30, wherein said antibody is a monoclonal antibody.

32. A method for detecting the presence of a BAG family protein in a sample, comprising the steps of:

- a. obtaining the sample;
- b. adding to said an antibody of claim 11 under suitable conditions for the binding of said antibody with the BAG family protein; and
- c. detecting said bound BAG family protein.

10 33. A method for detecting the presence of a first nucleic acid molecule that encodes a BAG family protein in a sample, comprising the steps of:

- a. obtaining the sample;
- b. adding to said sample a second nucleic acid molecule capable of hybridizing with said first nucleic acid molecule under suitable conditions for the binding of said second nucleic acid molecule with said first nucleic acid molecule; and
- c. detecting said hybridized first and second nucleic acid molecules.

20 25 34. A method of determining the risk of metastatic spread of cancer or prognosis of cancer patients by determining the level of expression of a BAG-family protein.

ACGGCGGCT CAGCTCCAT CGCTGGGG CGTACACAGTG CGGGCTGGC TCAGCCGGG 6666CCGGG GACCCGGGG CGACCCGGG 90
 L A Q R G G A R R P R G D R E
 BAG-1L

CGGCTGGTT CCCGGCTGG CGCCCTTCGG CCAGGGCGG AGCCGGGGCA GTCGGAGGCC CGGGCCCHGC GTGGTCCGCC TCCCTCTGG 180
 R L G S R L R A L R P G R E P R Q S E P A Q R G P P S R

CGTCCACCTG CCCGGAGTAC TGCCAGCGG CATGCCGAC CGACCCGGG CGCCGGCC GGGCTCGCA GGGCGGGAT GAGGAGAAA 270
 R P P A R S T A S G H D R P T R G A A G A R R P R H K K BAG-1M

ACCGGGGCC GCTCGAGCC GGGGGGGG TTGACCCGGG GCGGGGGGTT GACCTGAGT GAGGAGGGCA CCTGGAGCA AGAGGGGCAAC 360
 T R R S T R S E E L T R S E E L T L S E E A T W S E E A T

CAGAGTGGG AGGGAGCCCA GGGGAGAGAG ATGATCGGA GCCAGGGGGT GACCCGGGAC GAGGAGTCA CCCGGAGGCA CGAGGGTCAAC 450
 Q S E E A T Q G E E H N R S Q E V T R D E E S T R S E E U T
 BAG-1

AGGGAGGAA AGGGAGCCCA TGGGGCGGC TGGGCTCAC CCGACTGTCA CGGAGGCAAGGCA TGAGGAGGCA GACCTTCATG TTACCTCCCA GCAGGGCAAC 540
 R E E H A A G L T V T H S N E K H D L H U T S Q Q G S

FIG. 1 AGTGAACAG TTGTCAGA CCTGGCCAG GTTGGTCAAG AGGTCAATGG GGTTCACAG TCTTTCAAG AACTATATT TAACTGAA 630
 S E P V V Q D L A Q V V E E V V E V I G V P Q S F Q K L I F K G K

TCTCTGAGG AGATGGAAAC ACCGGTTGCA GCACTTGGAA TACAGATGG TTGCGGGGTC ATGTTAATTC GGAAAGAAA CAGTCACAG 720
 S L K E H E T P L S A L G I Q D G C R V M L I G K K H S P Q

GAGAGGGTTG AACTAAAGAA GTTGAACAT TTGGAGAAGT CTGTGGAGAA GATAGCTGAC CAGCTGGAAAG AGTGGATAAA AGAGCTTACT 810
 E E V E L K K L K H L E K S V E K I A D Q L E E L N K E L T

GGAATCCAGC AGGGTTCTT GCCCAGGAT TTGCAAGCTG AAGCTCTGT CAAACTGTAG AGGAGGATAA AGGCCACAAT AGGGAGTTT 900
 G I Q Q G F L P K D L Q A E A L C K L D R R U K A T I E Q F

ATGAGATCT TGGGGGAGAT TGACACACTG ATCCGGCGAG AARATTCA AGACAGTAGA TTGAAAGGA AGGGCTTGGT AAAAAAGGTT 990
 H K I L E E I D T L I L P E N F K D S R L K R K G L U K K V

CGGGATTCC TAGCGAGTG TGACACAGTG GAGGAGACAA TCTGCCAGGA GACTGAGGG CTGCACTGTA CAAACTTGC CCTGGCCGAG 1080
 Q A F L A E C D T V E Q N I C Q E T E R L Q S T N F A L A E

TGAGGTGTAG CAGAAAGGG CTGTGCTGCC CTGAGAGATG GCGCCACAG CTCTGCCGTC TCTGGATGG AATTACCTG ATTCTCTCAG 1170

GGCTGGGG GGCACACTGGC CATTGGCAA TTTTCCTACT CTCACACTGG TTCTCAATCA AAGATAGTGT CTTTGTGATT TGAGTAAGC 1260

TCCTTATCTG TTTTCAACAA AAAAAAAA A 1291

1/39

1

90

GCAGCCGGG TGTGGGAAAG TCTCTCCGGG TTGCCCCCGG GGGGTCAAGAG GGAGGGCGGG CGGCCGCTTG GTGACGGCGA CCCTGCAGGCC
CAGGGCCGC TCCACTGGCT GCGCCGGAG GGGCCGGTCA CTCCTGACTA CCCCGCTCG GAGGCTTGAAG TGGCTCAGGGC GAGGATCAAC
M A Q A K I N

GCTTAAAGCCA ACCTGGGGCG CTCTGCCGC TCCCTCCCA TGGCTGACCC CTCCTGGCGC CTGCTGGAGA GCCTGGACCA GCTGGAGGCTC
A K A N E G R F C R S S M A D R S S R L L E S L D Q L E L

AGGGTTGAGC CTTTGAGACA ACCGGCAACT GCTGTTGAGC AAGAGAAGAAGA AATCCTCTG GAAATGATCC ACAGTATCCA AATAGGCCAG
R V E A L R E A A T A V E Q E K E I L L E M I H S I Q N S Q

GACATGGGC AGATCAGTCG CGGAGGAAAGA GAGGAAATTAATCTGACTGC AAACCGTTG ATGGGAGAGA CTCTCACCCG TGAAGTGTCA
D H R Q I S D G E R E E L N L T A H R L H G R T L T V E U S

FIG. 2A V E T I R N P Q Q E S L K H A T R I I D E V U N K F L D D

TTGGGAAATG CCAGGAGTC TTAAATGTCG CTCTACAGTG CTGTTCATC TGAGGGCCA CATGGGCCAG TTGATCAGA GTTTCAATCC
L G N A K S H L M S L Y S A C S S E V P H G P V D Q K F Q S

ATAGTAAATTG GCTGTGTCT TGAGGATCAG AGGAAATTA AGAGGAGATT AGAGACTCTG CTAGAATAA TTGAAAGACTC TGACAGGCC
I V I G C A L E D Q K K I K R R L E T L R N I E M S D K A

ATCAGGCTAT TAGGGCATT TAGGGAGCT GGTCCAAAATCTGCAACAA AATGCTGAA AGCAGAGATTCA ATTAGTCCTTC AACCTAAGA
I K L L E H S K G R G S K T L Q Q N A E S R F N

2/39

630

720

810

3/39

GCATTTACAC AATACACAG GTGTAATAAT GATAAAATAC TATTTTAATT GATAACTAGT TCTTTGTTAG GATAACAC TTAGTTGACA 900
CTGATAGTTG TTTAGATGA GGAAATAATT CCATCAGTA TCTTCAGTT TGTAATAAC AATACAGTA ATATTTTAAT TATCTATCA 990
GAGATTTTT AGATTGATT CTTGCTGT ACTAGGATCT AGCATATTTC ACTATTCGTT GGATGAATAC ATAGTTGTC GGGAAACAA 1080
ACGTTAGCT AGGGCCAAA ACCATGATG CTTGGCTCTG TCTGGCATGG ATACCCCGAG TCAACCTGGG CATTAGTTT ACTAGGATT 1170
CTTACTGG 1179

FIG. 2B

4/39

GCGGAGCTCC GCATCCARACC CCGGGCCCGG GCCAAGTCTCT CTGGACTGGA CCAGAAGTTT CTAGCCGGCC AGTTGCTACC TCCCTTATC 90
 A E L R I Q P R A A A N F S G L D Q K F L A G O L L P P F I

 TCCCTCTTCC CCTCTGGCGAG CGAGGAGGCT ATTTCCAGAC ACTTCCACCC CTCTCTGGCC ACGTCACCCC CGCCTTAAAT TCATARRAGGT 180
 S S F P S G S E E A I S R H F H P S L A T S P P P L I H K G

 GCGGGCGCCG GGCTTCCCGG ACACGTGGC GGCGGAGGG GGCCACCGGC GGCGGGCCGG CGAGAGACTC GGCGCCGGA GCCAGCGCCC 270
 A R R R L P G H U G G G E G P T A A A R P E T R R P E P A P

 CGCACCCGGC CCCCAGCGGG CAGACCCAA CCCAGCATGA GCGCCGCCAC CCACCTCGCC ATGATGCGG TGGCGTCCGG CACCGTGA 360
 R T R A P A G R P Q P S M S A A T H S P M M Q V A S G N G D

 CGCGACCCCT TGCCCCCGG ATGGGAGATC AAGATCGACC CGCAGACCGG CTGGCCCTC TTCGTGGACC ACAAAGCCG CACCAACTACG 450
 R D P L P P G H E I K I D P Q T G H P F F U D H N S R T T T

 TGGGAGCACC CGCGCGTGCCT CTCTGAGGGC CCCAGGGAGA CTCCATCTC TGCCAAATGGC CCTTCCCGG AGGGCTCTAG GCTGCCGCT 540
 W N D P R U P S E G P K E T P S S A N G P S R E G S R L P P

 GCTAGGGAGG GCAACCCCTGT GTACCCCCAG CTCCGACCGG GCTACATTC CATTCTGTG CTCCATGAG GCGCTGAGA CCGGAGGTG 630
 A R E G H P V Y P Q L R P G Y I P I P U L H E G A E N R Q V

 CACCCCTTCC ATGTCTATCC CCAGCCTGGG ATGAGCGAT TCCGAACTGA GGCGCAGCA GCGGCTCTC AGAGGTCCCA GTACACCTCTG 720
 H P F H U Y P Q P G M Q R F R T E A A A A A P Q R S Q S P L

 CGGGGATGC CAGAACCCAC TCAGCCAGAT AAGCTGTG GAGCGGTGGC AGCGCGGGG GAGGCCAGC CCCAGCTCTC CCACCGACCT 810
 R G M P E T T Q P D K Q C G Q U A A A A A A Q P P A S H G P

 GAGCGGTCCC AGTCTCCAGC TGCCCTCTGC TGCTCTCTC CATCCTCTC GGCCAGCTG CCTTCCCTCG GAGGGAGCAG CCTGGCGAGT 900
 E R S Q S P A A S D C S S S S S S A S L P S S G R S S L G S

 CACCGCTCC CGCGGGGGTA CATCTCCATT CGCGTGATAC AGCGACAGAA CGTTACCCGG CGAGCAGCCC AGCCCTCTT CCACRAAGCC 990
 H Q L P R G Y I S I P V I H E Q N U T R P A A Q P S F H K A

 CAGAAGACGC ACTACCCAGC CGAGAGGGGT GAGTACCCAGA CCCACCAAGCC TGTGTACCCAC AGATCCAGG GGGATGACTG GGAGCCCCGG 1080
 Q K T H Y P A Q R G E Y Q T H Q P U Y H K I Q G D D W E P R

 CCCCTGGGG CGGCATCCCC GTTCAGGTCA TCTGTCAGG GTGCATCGAG CGGGGAGGGC TCACCGAGCA GGAGCAGCAC GCACTCCAC 1170
 P L R A A S P F R S S U Q G A S S R E G S P A R S S T P L H

 TCCCCCTCGC CCATCCCGTGT GCACACCGTG GTCGACAGGC CTCAGCAGCC CATGACCCAT CGAGAAACTG CACCTGTTTC CGACGCTGAR 1260
 S P S P I R U H T V U D R P Q Q P M T H R E T A P U S Q P E

 AACAAACCAAG AAGTAAGCC AGGCCAGTT GGACCAAGAC TCCCTCTGG ACACATCCCA ATTCAAGTGA TCCGCAAGA GGTTGATTCT 1350
 N K P E S K P G P U G P E L P P G H I P I Q V I R K E V O S

 AACACCTGTTT CCCAGAGGCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA AGTTCCCTCT GCTCCAGTTC CTTGCTCTCC TCCCAGCCCT 1440
 K P U S Q K P P P P S E K V E U K U P P A P U P C P P P S P

GGCCCTTCTG CTGTCCTCTC TTCCCCCAG AGTGTTGCTA CAGAAGAGAG GGCAAGCCCC AGCACTGCCC CTGCAGAAGC TACACCTCCA 1530
 G P S A U P S S P K S V A T E E R A A P S T A P A E A T P P

 AACCCAGGAG AAGCCGAGGGC TCCCCCAGGAGG CATCCAGGAG TGCTGAAAGT GGAGCCATC CTGGAGAAGG TGCGAGGGCT GGAGCAGGGCT 1620
 K P G E A E A P P K H P G V L K V E A I L E K V Q G L E Q A

 GTAGACAACCT TTGAGGGCAA GAGAGCTGAC AAAAGTACCC TGATGATCGA AGAGTATTG ACCAAGAGGC TGCTGGCCCT GGATTCACTG 1710
 V O N F E G K K T D K K Y L M I E E Y L T K E L L A L D S V

 GACCCCGAGG GACGAGGCCGA TGTGCGTCAG GCCAGGAGAG ACGGTTGAG GAGGTTAG ACCATCTGG AAAACTGCA ACAGAAGGCC 1800
 D P E G R A D U R Q A R R D G U R K V Q T I L E K L E Q K A

 ATTGTGTCC CAGGTCAAGT CCAGGTCTAT GAACTCCAGC CCAGCAACCT TGAGGCAGAT CAGCCACTGC AGGCAATCAT GGAGATGGGT 1890
 I D U P G Q V Q U Y E L Q P S N L E A D Q P L Q A I M E M G

 GCGCTGGAGG CAGACARAGGG CAGAAAGAT GCTGGAATG CAGAAGATCC CCACACAGAA ACCCAGCAGC CAGAGGCCAC AGCAGCAGCG 1980
 A U A A D K G K K N A G N A E D P H T E T Q Q P E A T A A A

 ACTTCACCAACC CCAGCAGCAT GACAGACACC CCTGGTAACC CAGCAGCAGC GTAGCCTCTG CCCTGTAAA GTCAGACTCG GAGCCAGATGT 2070
 T S M P S S M T D T P G N P A A P

GTGCTTTAGG GATTTAGTT GAGTCATTT CAGAGACTTT AGGTCAAGTT GTTTGATTA GTCGCTGGT ATGCAAGTACT TGTTGAGGGC 2160
 AACACATATA AAGGGCTAAA AGGGAAATG ATGCTTTCTC TCAATATTCT TACTCTGTG AATTAANGA AGTTGCTGT TTGTTGAGAA 2250
 GTTTAACCCC GTTGCTTGTT CTGCAGCCCT GTCACTTGG GCACCCAC CACCTGTTAG CTGTTGTTGT GCACTGTCTT TTGTACTCT 2340
 GGACTGGAGG GTAGATGGG GAGTCATTA CCCATCACAT AATATGAA CATTATCAG AATATGTCAG ATTTAATGAA GATGATTTTC 2430
 TTCATCTCAT AATTAAATA CCTGACTTTA GAGAGAGTAA AATGTGCCAG GAGCCATAGG AATATCTGTA TGTTGGATGTA CTTAATGCT 2520
 ACATTTTH FIG.3

AGATATCCT GTRAGACCA AAGTTGCAAG GCCAGAGTTT GAAATTCTAT AACAAATGGAG CGTATGGTCC AACATACCCC CCAGGGCCCTG 90
 GGGCAATAC TGCTCTATC TCGGGGCTT ATTATGCAAC TGGTTATACT CAGACCAAGTT ACTCCACAGA AGTTCAGAT ACTTACCGTT 180
 CATCTGGCAAA CAGCCCAACT CCAAGTCTCTC GTTGGATCTA TCCCAGAGG GRCGTCAAG ACTGAGACAC CCCCTCTTAA GGGGAGGTT 270
 CCAGGATATC CGCCTTCACA GAAACCTGGAA ATGACCCCTGC CCCATTATCC TTATGGAGAT GGTAATGCTA GTGTTCCACA ATCACGGCCG 360
 M E M V I V V F H N H G R
 ACTGTACGAC CACRAGGAG ATGGCTGGGC TTCTCCTGGT GCTTATGGAA TGGGTGGCG TTATCCCTGG CCTTCATCAG CGCCCTCAGC 450
 L Y D H K K D A W A S P G A Y G M G G R Y P W P S S A P S A
 ACCACCCGGC AATCTCTACA TGAATGAAAG TACCTCACCA TGGCCTAGCA GTGGCTCTCC CCGAGTCACCC CCTTCACCCC CAGTCAGGCA 540
 P P G N L Y M T E S T S P W P S S G S P Q S P P S P V Q Q
 GCCCAAGGAT TCTTCATACC CCTATAAGCCA ATCAGATCAA AGCATGAAAC GGCACAACTT TCCTTGCAAGT GTCATCAGT AGAAATCCTC 630
 P K D S S Y P Y S Q S D Q S M N R H N F P C S V H Q Y E S S
 GGGGACAGTG AACATGATG ATTCAAGATCT TTGGGATTC CAGTCAGT ATATGCTGA GCCTCAGCTG TATGGTAATG CCACCAATGTA 720 5/39
 G T V N N D D S D L L D S Q V Q Y S A E P Q L Y G N A T S D
 CCATCCCAAC AACATGAGTC AACATGAGAG TCTTCCTGAA GAAATGCTAC CTTCAAGATGA AAGTACTCTT CGGAGTATTA AACAAATCAT 810
 H P N N Q D Q S S L P E E C V P S D E S T P P S I K I
 ACATGTCTG GAGAAGGGTCC AGTATCTGA AACAAAGAGTA GAGGAATTG TAGGAAAGAA GAGAGACAAA GCATACCTGGC TTCTGGGAGGA 900
 H V L E K V Q Y L E Q E V E E F U G K K T D K A Y W L L E E
 AATGCTAACC AACGAACTTT TGGAACTGGAA ACTGGGGGUC AGGAACTGT AGGGAGGCC AGHAAAGGGG CTGTTTGTA 990
 M L T K E L L E L D S V E T G G Q D S V R Q A R K E A V C K
 GATTCAAGGCC ATATTGGAAA
 I Q A I L E 1010

FIG. 4

GAGGAAATAAA AATGAACTT CTCCAAAGCAC AARACCCCTTC TGAATTGTCAC CTGAGGCTCCA AARACGAGATT GCAGGGTTTA ATTGGACAGT 90
 E I K N E L L Q A Q N P S E L Y L S S K T E L Q G L I G Q L
 TGGATGAGGT AAGTNTGAA AAAAAACCCCT GCATCCGGGA AGCCAGGAGA AGAGCAGGAGA TCGAGGTGA AACTCTGTCGA AACTCTGTC ACATATATG 180
 D E U S X E K N P C I R E A R R R A U I E V Q T L I T Y I D
 ACTTGAGGGA GGCCCCCTGAG AAAAAAGAGC TGTTTGCTTG TGGGGAGCAC CCATCCCATA AAGCCGTCCTG GAAAGCTCTT GAAAGCTGT 270
 L K E A L E K R K L F A C E E H P S H K A V W H U L G N L S
 CTGAGGATCCA GGGAGGAGTT CTTTCATTG ATGGAAATCG AACCCTATAG AACTACATCC GGCTGGAGAGA GCTGCTCACC AAGCAGCTGC 360
 E I Q G E U L S F D G N R T D K N Y I R L E E L L T K Q L L
 TAGCCCTGGA TGCTGTTGAT CCGCAGGGAG AGAGGAGGT TAAAGGCTGCC AGGAAACAGG CTGTGAGGGCT TGCGCAGGAT ATTCTCAGCT 450
 A L D A U D P Q G E E K C K A A R K Q A V R L A Q N I L S Y
 FIG.5 ATCTCGACCT GAAATCTGAT GAAATGGGAGT ACTGAAATAC CAGAGATCTC ACTTTGATA CTGTTTGCATCTATATGT GCTTCATATGT 540
 L D L K S D E W E Y 6/39
 ATAGAGAGCT TTCAAGTTCAT TGAATTATAC GTGCAATTGTT CAGTCTCAGT ATTTATGATT GAAAGCAGATT CTATTCAAGTA TCTGCTGCCTT 630
 TTGATGTTGC AAGACAAATA TCATTACAGC ACGTTTACCTT TTCCATTGGG ATCAAAAA 689

7/39

ATGTCTTCCGCCTTCGTTGAAATATTCACTTCTTTCCAGCTTTCCCCATCTGAC
CT
GCTTTGGTTTT
CGAGAAAACACGTTCAAATCAGCGACATCTCTCAAATTGAGATCATAGGCTTTGAAGA
TTGCTCAAATTATG
CTTCTCATATTGCATGAGCATTGAAAGCCCGCGTCATCAACCAAAGCATTTCACCCAT
CACAATGATTATCATTCTTAAATT

FIG. 6A

8/39

MKVNVSCSSV	QTTIDILEEN	QGEDESIITL	GQLRDRIATD	NDVDVETMKL	50
LHRGKFLQGA	DDVSLSTLNF	KENDKIIVMG	GKNALVDDAG	FKMLMQYEKH	100
NLSNLQKAYD	LNLRDVADEL	RGFLEKPKQV	EMGKKLEKKV	KYFNEEAERH	150
LETLDGMNII	TETTPENQAK	RNREKRKTLV	NGIQTLLNQN	DALLRRLQEY	200
QSVLNGDIPE					210

FIG. 6B

9/39

ATGCCAGTCG	TGAACATACC	AATCAAATA	CTTGGTCAGA	ATCAATCACA	50
TAGTCGAAGT	AACTCCTCGT	CTTCTGTTGA	CAACGATCGA	AATCAACCAC	100
CACAGCAGCC	ACCTCAACCG	CAACCACAAC	AGCAATCTCA	GCAACAATAC	150
CAGCAGGCTC	CAAACGTGAA	TACCAATATG	CATCATTCCA	ACGGATTCTC	200
ACCTAACTTC	CCATCTCGTA	GTCCTATTCC	GGACTTTCCC	AGTTTTTCAT	250
CTGGGTTCCC	AAACGATTCT	GAATGGTCTT	CGAATTCCC	GTCGTTCCA	300
AATTTCCTAA	GTGGATTCTC	AAATGGAAGT	TCTAATTTC	CTGATTTCC	350
AAGATTCGGA	AGAGATGGAG	GAATCTGCC	AAACCCACCG	ATGCAAGGAT	400
ACAGGAGAAG	TCCAACACCA	ACATCAACTC	AATCTCCAAC	TTCTACATTA	450
AGACGCAACT	CTCAGCAGAA	TCAAGCTCCT	CCACAATATT	CTCAGCAACA	500
ACCACAACAA	GCTCAACAAC	GTCAGACAAAC	TCCTCCGTCA	ACAAAAGCTT	550
CATCTCGACC	ACCATCTCGT	ACTCGTGAAC	CAAAGGAACC	TGAGGTACCC	600
GAGAGACCAG	CAGTTATTCC	ATTGCCATAT	GAGAAGAAGG	AGAAAACACT	650
GGAGAAGAAA	GGTAGTCGTG	ATTCTGGAAA	GGGTGATGAG	AACCTTGAAG	700
AGAACATTGC	CAAGATCACG	ATCGGAAAGA	ATAATTGCAGA	GTTATGTCCG	750
GAACAAGAAA	CGGACGGCGA	CCCATCTCCA	CTAACCTCCC	CAATCACCGA	800
AGGAAAGCCA	AAGAGAGGAA	AGAAAACCTCA	ACGTAATCAA	AGTGTGTTG	850
ATTTCAATGC	CAAGACAATT	GTTACTTGG	ATAAAATTGA	ATTACAAGTT	900
GAGCAGTTGA	GAAAAAAAGC	TGCTGAACTC	GAAATGGAAA	AAGAGCAAAT	950
TCTTCGTTCT	CTAGGAGAAA	TCAGTGTCA	TAACTGCATG	TTCAAACATGG	1000
AAGAATGTGA	TCGTGAAGAG	ATTGAAGCAA	TCACTGACCG	ATTGACAAAA	1050
AGAACAAAGA	CAGTTCAAGT	TGTTGTCGAA	ACTCCACGAA	ATGAAGAACAA	1100
GAAAAAAAGCA	CTGGAAGATG	CAACTTTGAT	GATCGATGAA	GTCGGAGAAA	1150
TGATGCATTC	GAATATTGAA	AAGGCTAACG	TGTGCCTACA	AACCTACATG	1200
AACGCCCTGTT	CGTACGAAGA	AACTGCTGGA	GCCACCTGCC	AAAACCTCTT	1250
GAAGATCATA	ATTCAAGTGC	CTGCTGATGA	TCAGAAACGC	ATCAAGCGTC	1300
GTCTGGAAAA	TCTGATGTCT	CAAATTGAGA	ATGCTGAGAG	AACGAAAGCA	1350
GATTTGATGG	ATGATCAAAG	CGAATAG			1377

FIG. 7A

10/39

MPVVNIPIKI	LGQNQSHSRS	NSSSVNDR	NQPPQQPPQP	QPQQQSQQQY	50
QQAPNVNTNM	HHSNGFSPNF	PSRSPIDFP	SFSSGFPNDS	EWSSNFPSPF	100
NFPGFSNGS	SNFPDFPRFG	RDGGLSPNPP	MQGYRRSPTP	TSTQSPTSTL	150
RRNSQQNQAP	PQYSQQQPQQ	AQQRQTTPPS	TKASSRPPSR	TREPKEPEVP	200
ERPAVIPLPY	EKKEKPLEKK	GSRDSGKGDE	NLEENIAKIT	IGKNNCELCP	250
EQETDGDPS	LTSPITEGKP	KRGKKLQRNQ	SVVDFNAKTI	VTLDKIELQV	300
EQLRKKAAEL	EMEKEQILRS	LGEISVHNCM	FKLEECDREE	IEAITDRLTK	350
RTKTVQWVE	TPRNEEQKKA	LEDATLMIDE	VGEMMHSNIE	KAKLCLQTYM	400
NACSYEETAG	ATCQNFLKII	IQCAADDQKR	IKRRLENLMS	QIENAERTKA	450
DLMDQSE					458

FIG. 7B

11/39

ATGTCAGAAA	AGACTAGCAC	AGTTACAATA	CACTATGGAA	ATCAGCGATT	50
TCCGGTAGCA	GTCAATCTAA	ATGAGACGTT	AAGTGAAGTG	ATTGATGATT	100
TACTGAAAC	GAATGAGATT	TCTGAGAAGA	AAGTCAGCT	TTTTTACGCT	150
GGCAAGCGTT	TAAAAGACAA	AAAAGCCTCG	TTATCAAAT	TGGGTTTAAA	200
AAATCATAGT	AAAATTCTAT	GTATAAGACC	ACATAAGCAA	CAACGAGGTT	250
CCAAGGAAAA	AGACACGGTT	GAGCCCCTC	CGAAAGCGGA	AGCGGAGAAT	300
CCTGTATTTT	CGCGTATTTT	TGGAGAAATA	AAAGCCATCG	ATCAGTATGT	350
TGACAAAGAA	CTTTCCCCCA	TGTACGACAA	TTACGTAAAT	AAACCGTCGA	400
ACGATCCAAA	GCAGAAAAAC	AAACAGAAAC	TAATGATAAG	TGAACCTACTT	450
TTACAACAGC	TTTTAAAATT	GGATGGAGTT	GACGTACTGG	GCAGCGAGAA	500
ATTGCGTTTT	GAACGGAAGC	AACTTGTTC	TAAGATCCAA	AAAATGTTGG	550
ATCACGTTGA	CCAAACAAGC	CAAGAAGTGG	CCGCATAG		588

FIG. 8A

12/39

MSEKTSTVTI	HYGNQRFPVA	VNLNETLSEL	IDDLLETTEI	SEKKVLFYA	50
GKRLKDKKAS	LSKLGLKNHS	KILCIRPHKQ	QRGSKEKDTV	EPAPKAEAEEN	100
PVFSRISGEI	KAIDQYVDKE	LSPMYDNYVN	KPSNDPKQKN	KQKLMISELL	150
LQQLLKLDGV	DVLGSEKLRF	ERKQLVSKIQ	KMLDHVDQTS	QEVA	195

FIG. 8B

13/39

ATGTCTTTT	TTACCCAGTT	GTGTTCTATG	GATAAAAAT	ATTGGATCTC	50
TCTAGCTGTA	TTGTCAGTTA	CTGTTTGAT	TAGCGCATT	TTGAAAAAGA	100
GAGCTACTGA	AACCGAAGAT	ATTGTCGTTG	TTCATTACGA	TGGCGAAAAG	150
TTGAATTTG	TGTTGCGACA	ACCAAGGCTG	AATATGGTTT	CTTACACTAG	200
TTTTCTTCGT	CGCGTGTGCA	ACGCATTTTC	AGTAATGCC	GACAAAGCGT	250
CTCTCAAGTT	AAACGGGGTG	ACCCTCAAGG	ATGGTTCACT	TTCCGACCAA	300
AATGTGCAAA	ATGGAAGTGA	ATTAGAGCTC	GAATTACCA	AACTGAGCCC	350
GGCAATGCAA	CAAATTGAAG	CATATATAGA	TGAGCTTCAA	CAGGATCTCG	400
TCCCTAAAAT	TGAAGCCTTC	TGCCAATCGT	CTCCCGCTTC	GGCACAAAGAT	450
GTTCAAGATT	TGCATACACG	CCTTAGTGA	ACATTGTTGG	CTAGGATGAT	500
AAAATTAGAT	GCTGTTAATG	TTGAAGACGA	CCCAGAAGCT	CGTCTTAAAA	550
GAAAAGAAGC	TATTCGTTTA	TCTCAACAAT	ATTTGAGTAA	ACTAGATTCC	600
ACCAAGAAC	AAAACAAATG	A			621

FIG. 9A

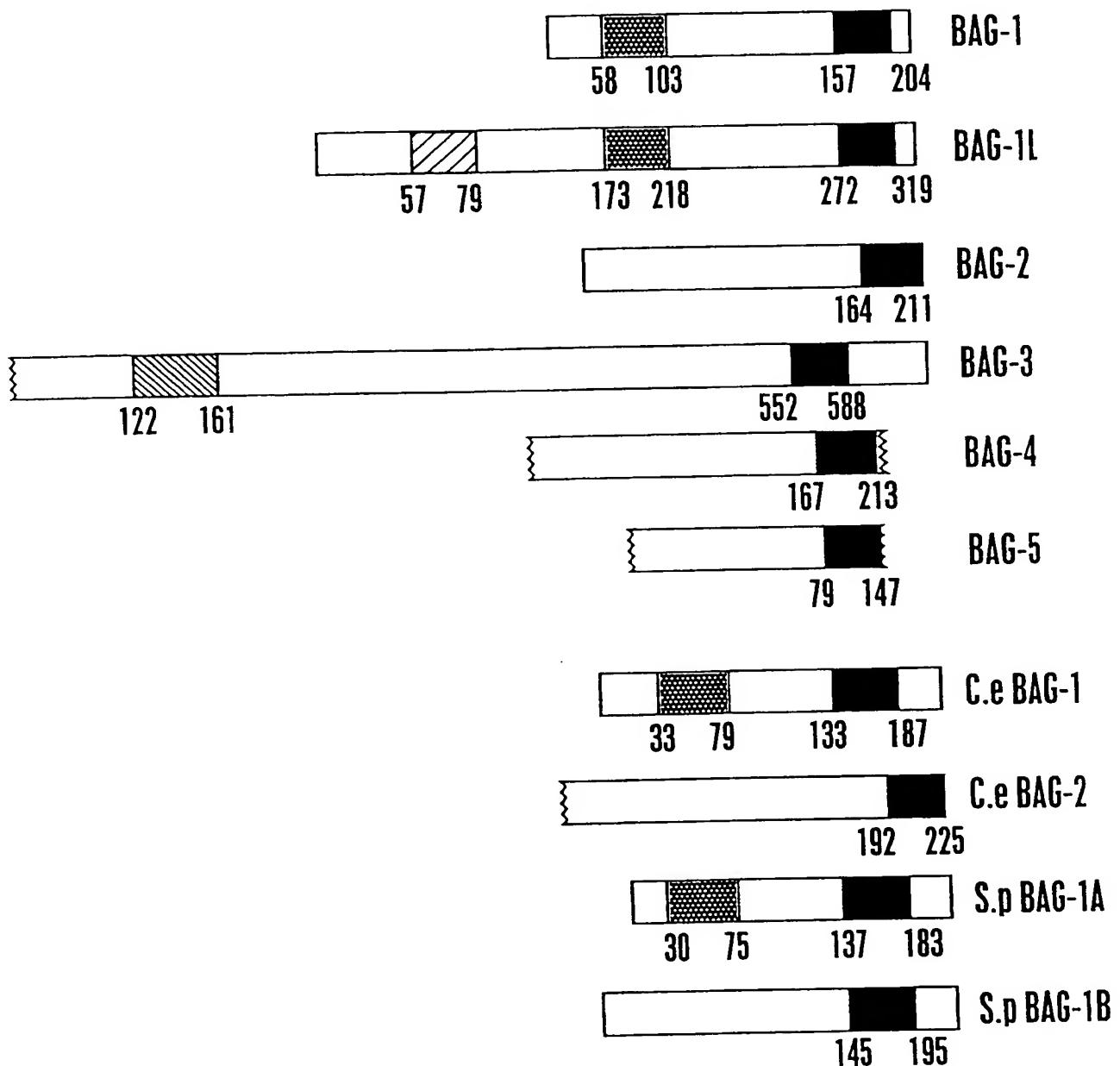
14/39

MSFFTQLCSM	DKKYWISLAV	LSVTVLISAL	LKKRATETED	IVVVHYDGEK	50
LNFVLRQPR	NMVSYTSFLR	RVCNAFSVMP	DKASLKNGV	TLKDGSLS	100
NVQNGSELEL	ELPKLSPAMQ	QIEAYIDELQ	QDLVPKIEAF	CQSSPASAQD	150
VQDLHTRLSE	TLLARMIKLD	AVNVEDDPEA	RLKRKEAIRL	SQQYLSKLD	200
TKNQNK					206

FIG. 9B

15/39

Fig.10A



■ Ubiquitin-Like

■ BAG Domain

■ WW

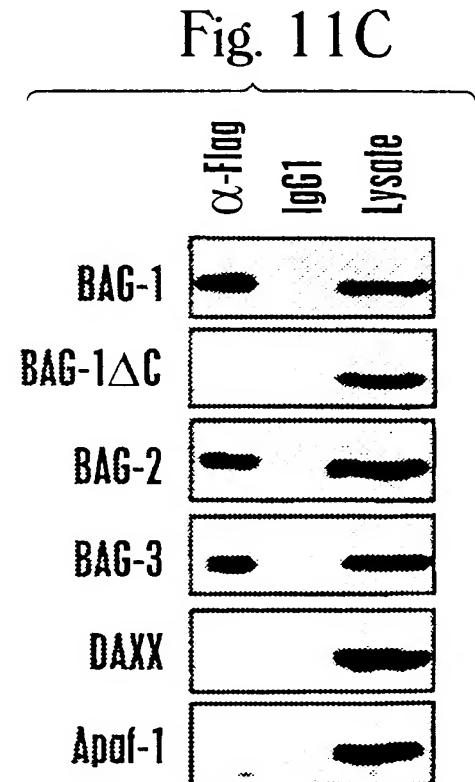
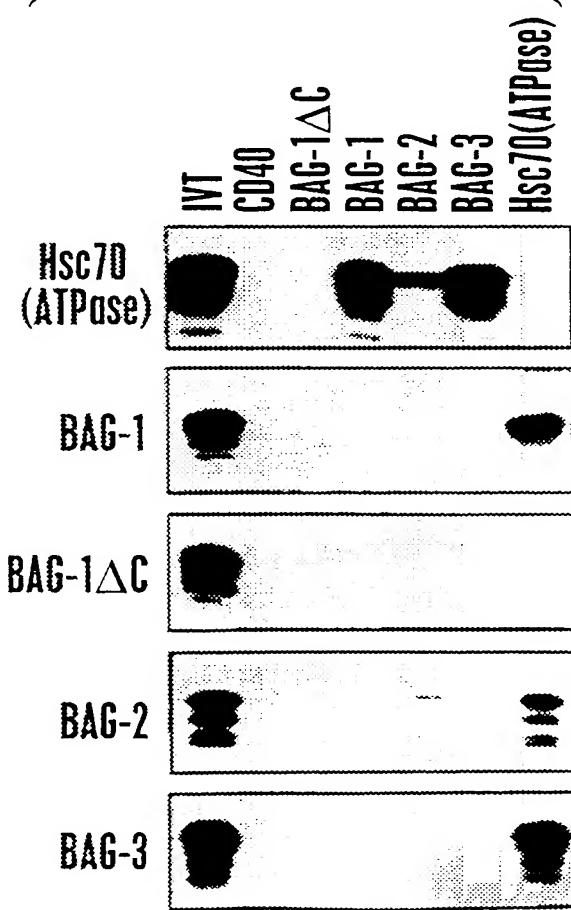
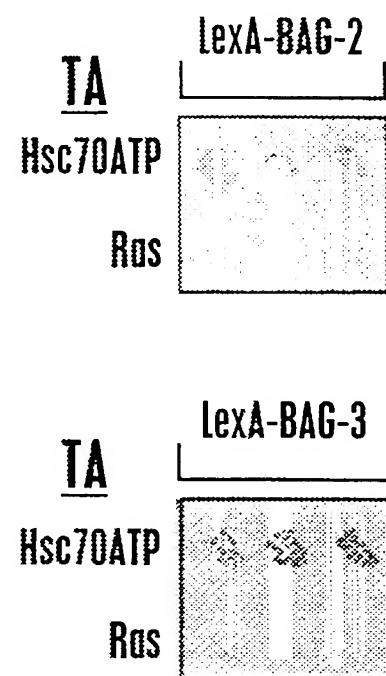
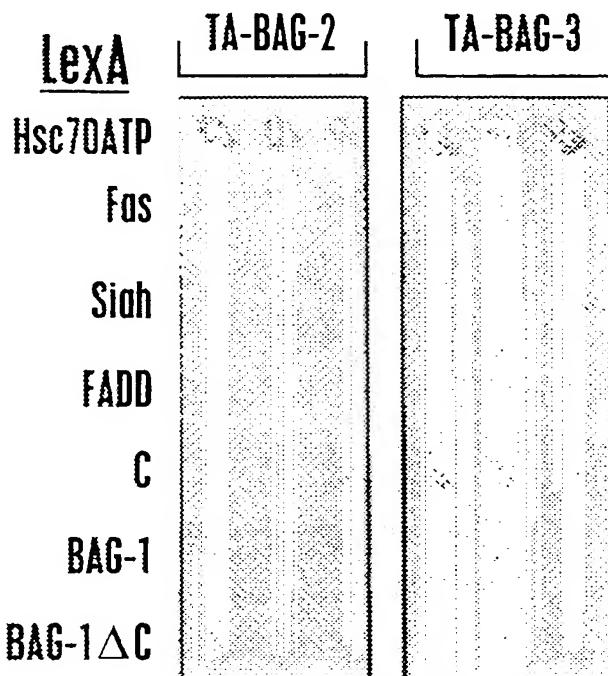
■ Nuclear Localization Signal

16/39

157	C KLD	R VKATI	E QF	R ILEE	I T	-	IP	E	-	-	-	N F K	D S	R L K R K	G	L V	Y	K V Q A F	I		
552	KK T	DKK YLM	EE Y	IT K	ELD S	VD	P E	G R A	-	-	-	-	D V R	Q A	R	G	V R K V Q	Y	I L		
167	KK T	DK A YWL	EE E	IT K	ELD S	V E	G	D	-	-	-	-	S	V R	Q A	R K E	A	V C	K		
79	KK T	DK Y YI	R L E	LL T	K Q L L	A	V D	P Q E	E	-	-	-	-	-	-	-	Q A V R L	Q	Y		
124	C KLD	R KV KATI	E QF	R ILEE	I T	-	V L	P	E	-	-	-	Q F K	D S	R L K R K	G	L V	X	K V Q	F	I
133	KKL	R KKV K YF	EE E	EE E	EE E	EE E	LDG	Y	LDG	Y	LDG	Y	R E	K R K	R E	K R K	Y	Y	Y	Y	
137	KK	R K K K	R K L M	R K L M	R K L M	R K L M	ELL	Q	OLL	K	LDG	V D V L	G	E	-	-	-	R F E	R K	R K	
145	DAY	R D L H T	R L S	R L S	R L S	R L S	ET	L L A	R I K	L D A	V	V	E	P	-	-	-	R L K R K	E	A	
164	LED	R K K	R K R R L	R K R R L	R K R R L	R K R R L	ET	L L	R M I	E	SI K	Y	I K	I L E H S K G A G S	K	Q Q	Y	A E	Y	Y	
192	ADD	R K R	R K R L	R K R L	R K R L	R K R L	ET	E	A E R	K	D L	-	-	-	-	-	-	R F N	Y	Y	

SUBSTITUTE SHEET (RULE 26)

Fig. 10B



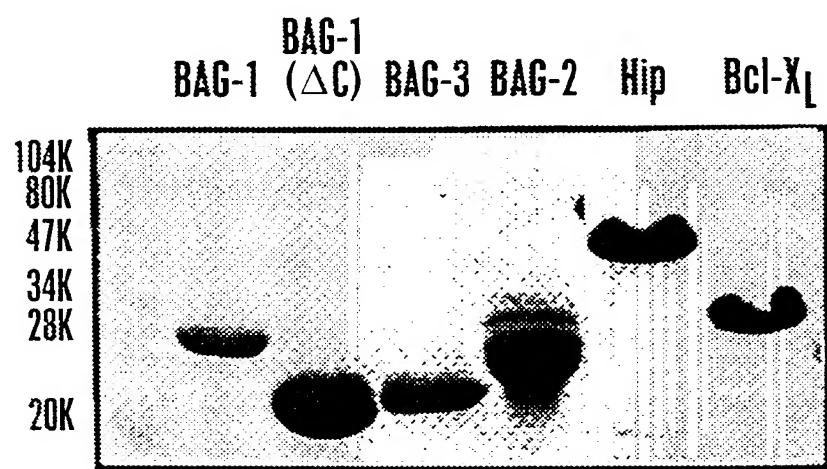
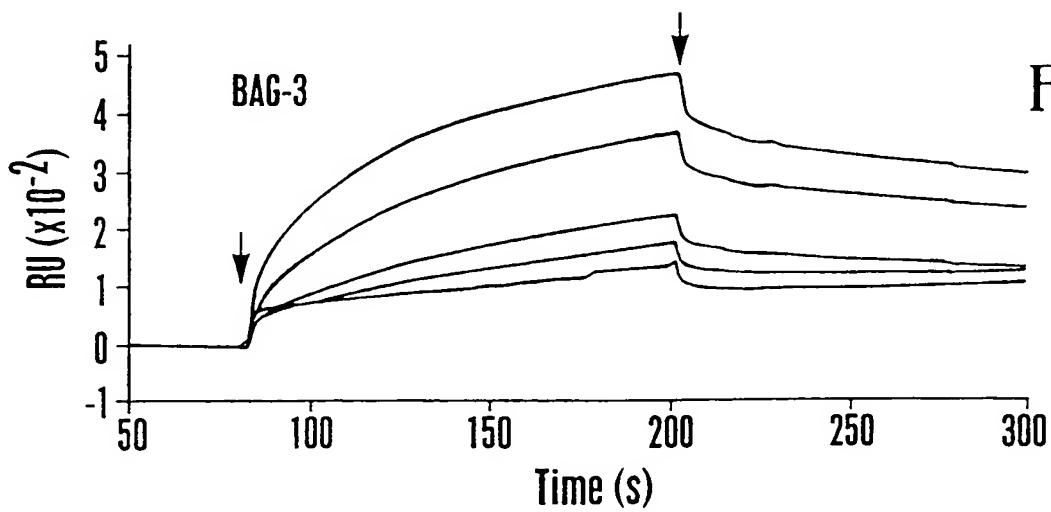
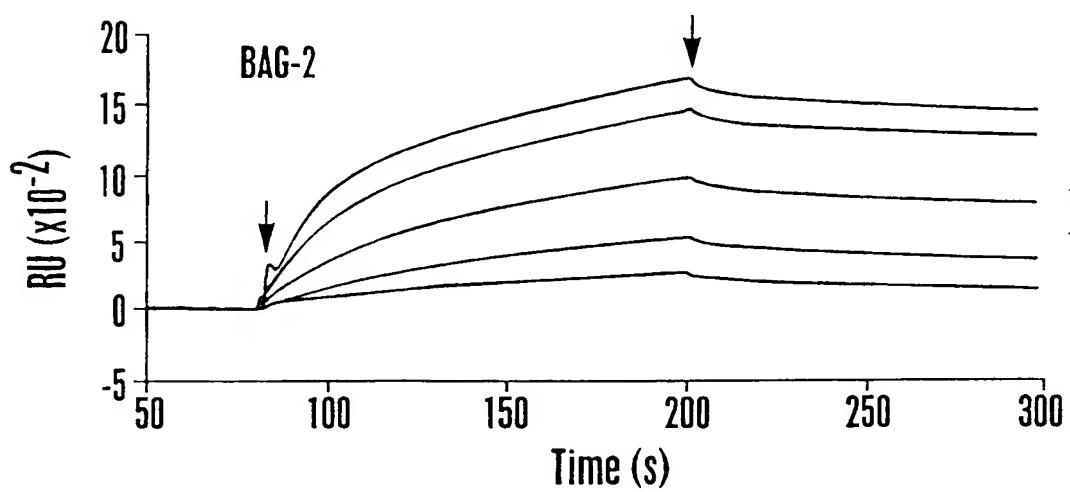
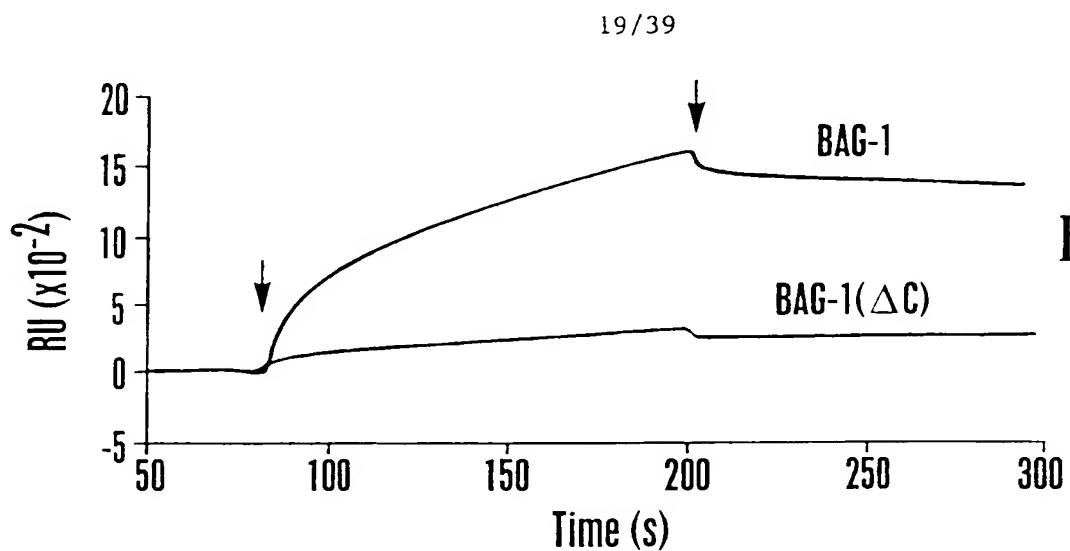
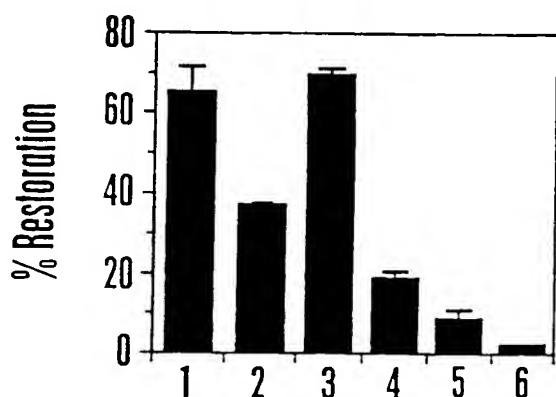


Fig. 12



20/39

Fig. 14A



	1	2	3	4	5	6
Hsc70	+	+	+	+	+	-
DnaJ	+	+	+	+	+	-
BAG-1	-	+	-	-	-	-
BAG Δ C	-	-	+	-	-	-
BAG-2	-	-	-	+	-	-
BAG-3	-	-	-	-	+	-
BSA	-	-	-	-	-	+

Fig. 14B

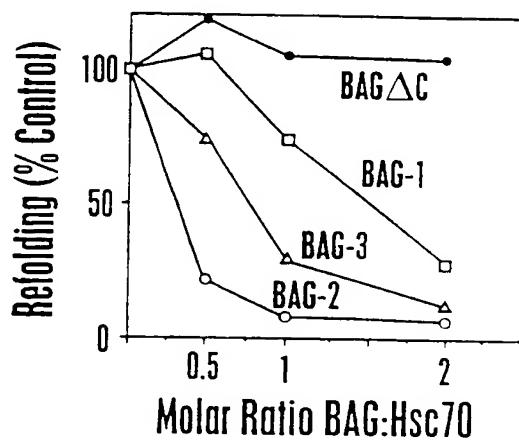
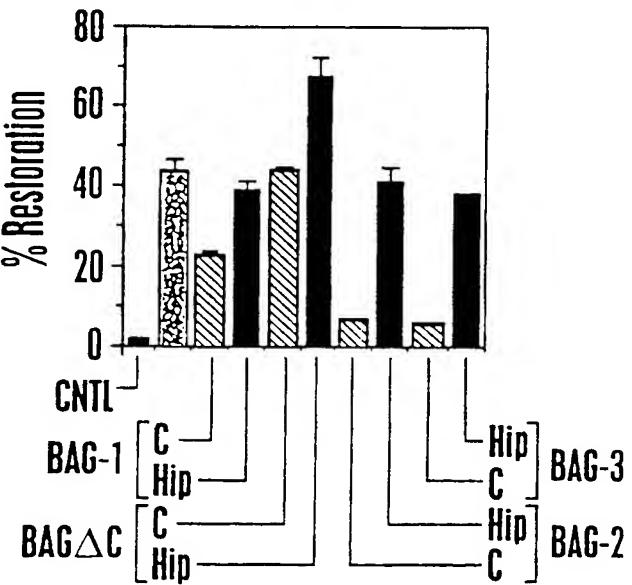


Fig. 14C



21/39

FIG. 15A

GGGGAGCTCC	GCATCCAAAC	CCGGGCCGCC	GCCAACTTCT	CTGGACTGGA	50
CCAGGAAGTTT	CTAGCCGGCC	AGTTGCTACC	TCCCCTTATC	TCCTCCTTCC	100
CCTCTGGCAG	CGAGGAGGCT	ATTTCCAGAC	ACTTCCACCC	CTCTCTGGCC	150
ACGTCACCCC	CGCCTTTAAT	TCATAAAGGT	GCCCCGCC	GGCTTCCCGG	200
ACAOGTGGC	GGGGAGAGG	GGOOCAOGGC	GGGGOOGG	CCAGAGACTC	250
GGGGGGAGA	GGCAOGGCC	GGCAOOOGCG	GGCAOGGGG	CAGACCCCCA	300
OCCAGCATGA	GOGOGGCCAC	CCACTOGGCC	ATGATGCAGG	TGGCGTCCGG	350
CAACGGTGA	CGCGGACCCCTT	TGCCCCCGGG	ATGGGAGATC	AAGATCGACC	400
CGCAGACGG	CTGGCCCTTC	TTCGTGGACC	ACAACAGGCC	CACCACTACG	450
TGAAACGACC	CGCGCGTGC	CTCTGGGGC	CCCCAAGGAGA	CTCCATCCTC	500
TGCAATGGC	CCTTCCCCGG	AGGGCTCTAG	GCTGCCGCGCT	GCTAGGGAAAG	550
GCCACCCCTGT	GTACCCCCAG	CTCCGACCAG	GCTACATTC	CATTCTCTGTG	600
CTCCATGAAG	GCCCCCTGAGAA	CCGGCAGGTG	CACCCCTTCC	ATGTCTATCC	650
CCAGGCTGGG	ATGCAAGGGAT	TCCGAACCTGA	GGGGGCAGCA	GGGGGCTCCTC	700
AGAGGTCCC	GTCACCTCTG	GGGGCATGC	CAGAAACCAC	TCAGGCCAGAT	750
AAACAGTGTG	GACAGGGTGGC	AGGGGGGGC	GCAGCCCCAGC	CCCCAGCCTC	800
CCACGGACCT	GAGGGTCCC	AGTCTCCAGC	TGCTCTGTGAC	TGCTCATCCT	850
CATCCTCTC	GGCCAGCC	CCTTCCTCG	GCAGGGCAG	CCTGGGCAGT	900
CACCAAGCTC	GGGGGGTA	CATCTCCATT	CCGGTATAAC	ACGAGCAGAA	950
CGTTACCGGG	CCAGCAGGCC	AGCCCTCTT	CCACAAAGCC	CAGAAAGACGC	1000
ACTACCCAGC	GCAGGGGGGT	GAGTACCCAGA	CCACACAGCC	TGTGTACCCAC	1050
AAGATCCAGG	GGGATGACTG	GGAGGCCCCGG	CCCCTGCGGG	CGGCATCCCC	1100
GTTCAGGTCA	TCTGTCCAGG	GTGCATCGAG	CCGGGAGGGC	TCACCCAGCCA	1150
GGAGCAGCAC	GCCACTCCAC	TCCCCCTCGC	CCATCGTGT	GCACACCGTG	1200
GTCCGACAGGGC	CTCAGCGGCC	CATGACCCAT	CGAGAAACTG	CACCTGTTTC	1250
CCAGCCTGAA	AACAAACCAAG	AAAGTAAGGCC	AGGGCCAGTT	GGACCCAGAAC	1300
TCCCCTGG	ACACATCCCA	ATTCAAGTGA	TCCGCAAAGA	GGTGGATTCT	1350

22/39

FIG. 15B

AAACCTGTT CCCAGAACCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA 1400
 AGTCCCCCT GCTCAGTT CTTGCTTCC TCCCAGCCCT GGCCTCTG 1450
 CTGTCCTC TTCCCCAAG AGTGTGGCTA CAGAAGAGAG GGCAGCCCCC 1500
 AGCACTGCC CTGAGAAC TACACCTCCA AAACCAAGGAG AAGCCGAGGC 1550
 TCCCCAAA CATCAGGAG TGCTGAAAGT GGAAGGCCATC CTGGAGAAGG 1600
 TGCAGGGCT GGAGCAGGCT GTAGACAACCT TTGAAGGCAA GAAGACTGAC 1650
 AAAAAGTACC TGATGATCGA AGAGTATTTG ACCAAAGAGC TGCTGGCCCT 1700
 GGATTCACTG GACCCGAGG GACGAGGCCGA TGTGCGTCAG GCCAGGAGAG 1750
 ACGGTGTCAG GAAGGGTCAG ACCATCTGG AAAAACCTTGAAACAGAAAGCC 1800
 ATTGATGTC CAGGTCAAGT CCAGGTCTAT GAACTCCAGC CCAGCAAACCT 1850
 TGAAAGCAT CAGCCACTGC AGGCAATCAT GGAGATGGGT GCCGTGGCAG 1900
 CAGACAAGGG CAAGAAAAT GCTGGAAATG CAGAAGATCC CCACACAGAA 1950
 ACCCAGCAGC CAGAAGGCCAC AGGAGCAGCG ACTTCAAAC CCAGCAGCAT 2000
 GACAGACACC CCTGGTAACC CAGCAGCACC GTAGCCTCTG CCCTGTAAAA 2050
 ATCAGACTCG GAACCGATGT GTGCTTTAGG GAATTAAAG TTGCATGCAT 2100
 TTCAAGACT TTAAAGTCAGT TGGTTTTAT TAGCTGCTTG GTATGCAGTA 2150
 ACTTGGGGGG AGGCAAAACA CTAATAAAAAG GGCTAAAG GAAAATGATG 2200
 CTTTCTTCT ATATTCTTAC TCTGTACAAA TAAAGAAGTT GCTTGTGTT 2250
 TGAGAAGTT AACCCCGTTG CTTGGTCTGC AGCCCTGTCT ACTTGGGCAC 2300
 CCCCACCC TGTAGCTGT GGTGTGCAC TGTCTTTGT AGCTCTGGAC 2350
 TGAGGGGGTA GATGGGGAGT CAATTACCCA TCACATAAAAT ATGAAACATT 2400
 TATCAGAAAT GTTGCATT TAATGAGATG ATTTTCTTCA TCTCATAATT 2450
 AAAATACCTG ACTTTAGAGA GAGTAAAATG TGCCAGGAGC CATAGGAATA 2500
 TCTGTATGTT GGATGACTTT AATGCTACAT TTTC 2534

23/39

FIG. 15C

MSAATHSPMM QVASGNGDRD PLPPGWEIKI DPQTGWPFV DHNSRTTTWN 50
DPRVPSEGPK ETPSSANGPS REGSRLPPAR EGHPVYPQLR PGYIPIVPLH 100
EGAENRQVHP FHVYPQPGM Q RFRTEAAAAA PQRSQSPLRG MPETTQPDKQ 150
CGQVAAAAAA QPPASHGPER SQSPAASDCS SSSSSASLPS SGRSSLGSHQ 200
LPRGYISIPV IHEQMNTRPA AQPSFHKAQK THYPAQRGEY QTHQPVYHKI 250
QGDDWEPRPL RAASPFRSSV QGASSREGSP ARSSTPLHSP SPIRVHTWD 300
RPQQQPMTHRE TAPVSQOPENK PESKPGPVGCP ELPPGHIPIQ VIRKEVDSKP 350
VSQKPPPSE KVEVKVPPAP VPCPPPSPGP SAVPSSPKSV ATEERAAPST 400
APAEATPPKP GEAEAPPKHP GVLKVEAILE KVQGLEQAVD NFEKGKTDKK 450
YLMIEEYLTK ELLALDSVDP EGRADVROQAR RDGVRKVQTI LEKLEQKAIID 500
VPGQVQVYEL QPSNLEADQP LQAIMEMGAV AADKGKKNAG NAEDPHTETQ 550
QPEATAAATS NPSSMTDTPG NPAAP 575

24/39

GGGAGCTCC GCATCCACC CCCGGCCGG GCAAACCTCT CTGACTGGA CAGAAAGTTT CTAGCCGGCC AGTTGCTAAC TCCCCTTTATC 90
 CCTCTTCCC CCTCTGGAG CGAGGAGGCT ATTCAGAC ACTTCCACCC CTCTCTGGCC CGCTAACCCCC CGCTAAAGGT 160
 GCGGGGGCC GCGTTCCGG ACACGTGGC GCGGGAGGG GCGGCAACGGG GCGGGGGGG CGAGAGACTC GCGGGGGGG GCGGGGGCC 270
 CGACCCGGG CGCCAGGGG AGACCCCAA CGACCCCAA CCCAGCATGA GCGGGCCAC CCACTGGCCC ATGATGAGG TGGCGTGGG CAAGGGTGA 360
 M S A A T H S P M Q V A S G N G D
 CGGACCCCTT TGCCCCGGG ATGGGAGATC AAGATGCCCG CGCAGACGCC CGGAGACCCCTT TTGGGACCC ACAACAGCCG CACCACTACG 450
 R D P L P P G W E I K I D P Q T G W P F F V D H N S R T T T
 TGGAAAGACC CGGGGGCC CTCTGAGGGC CCCAAGGAGA CTCATCCCTC TGCCAAATGGC CTTCCGGG AGGGCTCTAG GCTGGGGCT 540
 W N D P R V P S E G P K E T P S S A N G P S R E G S R L P P
 GCTAGGGAG GCCACCCCTGT GTACCCCCAG CTCCGACAG GCTACATTCC CATTCTGTG CTCATGAAG GCGCTGAGAA CGGGAGGTG 630
 A R E G H P V Y P Q L R P G Y I P V L H E G A E N R Q V
 ACCCCCTTC ATGCTATCC CGAGCCGGAT ATGAGGGAT TCCGAACCTGA GGGGAGGA GGGGCTCTC AGAGGCCCCA CTCACTCTG 720
 H P F H V Y P Q P G M Q R F R T E A A A A P Q R S Q S P L
 CGGGGATGC CAGAACCCAC TCAGCCAGAT AACAGGTGTG GACAGGTGGC AGGGGGGGC GCAAGCCCCAGC CCCCAGGCTC CCACGGACCT 810
 R G M P E T T Q P D K Q C G Q V A A A A A A Q P P A S H G P
 GAGGGCTCC AGTCTCCAGC TGCTCTGTAC TGCTCATCC CATCCCTCTC GGGCAGGCTG CTTCTCTCG GCAAGGGAG CCTGGGAGT 900
 E R S Q S P A A S D C S S S S S A S L P S S G R S S L G S
 CACAGCTCC CGGGGGGTA CATCTCCATT CGGGTGTATAC ACAGGAGAA CGTTACCCGG CCAGGAGGCC AGCCCTCCCT CCACAAAGCC 990
 H Q L P R G Y I S I P V I H E Q N V T R P A A Q P S F H K A
 CAGAAAGCC ACTACCCAGC GGAGGGGGT GAGTACCGA CCCACAGCC TGTGTACAC AAGATCCAGG GGAGCCCCGG 1080
 Q K T H Y P A Q R G E Y Q T H Q P V Y H K I Q G D D W E P R
 CCCCCCTGGG CGGCATCCCC GTCAAGGTCA TCTGTCAAGG GTGCATCGAG CGGGGAGGG TCACCAAGCC GCAACTCCAC 1170
 P L R A A S P F R S S V Q G A S S R E G S P A R S T P L H
 TCCCCCTGGC CGATCCGTGT GCAACCCGGC GTGGACAGGC CTAGGAGGCC CATGACCAT CGAGAAACTG CACCTGTTTC CGAGGCTGAA 1260
 S P S P I R V H T V V D R P Q Q P M T H R E T A P V S Q P E
 AACAAACAG AAAGTAAGCC AGCCCCAGTT GGACCGAAC TCCCTCTGG ACACATCCA ATTCAACTGA TCCCACAAAGA CCTGGATTCT 1350
 N K P E S K P G P V G P E L P P G H I P I Q V I R K E V D S

Fig. 15D

25/39

AACCTGTT CCCAGAAGCC CCCACCTCCC TCTGAGAAGG TAGAGGTGAA AGTTCCCCCT GCTCAGTTC CTTGTCCTTC TCCCCAGCCCC 1440
 K P V S Q K P P S E K V V P V P A P V P C P P P S P
 GCCCCCTCTG CTGCTCTCTG TTCCCCAAG AGTGGCTA CAGAAGAGAG GGCAGCCCCC AGCACCTGCC CTCAGAAGC TACACCTCCA 1530
 G P S A V P S S P K S V A T E E R A A P S T A P A E A T P P

 AACCCAGGAG AAGCCGAGGC TCCCCAAAAA CATCCAGGAG TGGTCAAAGT GGAAGCCATC CTGGAGAAGG TGGAGGGCT GGAGGAGGGCT 1620
 K P G E A E A P P K H P G V L K V E A I L E K V Q G L E Q A

 GTAGACAAT TTGAAAGCAA GAAAGACTGAC AAAAGTACCG TGATGATCGA AGAGTATTG ACCAAAGAGC TGCTGGCCCT GGATTCACTG 1710
 V D N F E G K K T D K K Y L M I E E Y L T K E L L A L D S V

 GACCCGGG GACGGGGGA TGTCGTCAG GCCAGGGAG ACCGGTTAG ACCATCTGG AAAAACTTGA ACAGAAAGCC 1800
 D P E G R A D V R Q A R R D G V R K V Q T I L E K L E Q K A

 ATGATCTC CAGGTCAGT CCAGCTCAT GAACTCCAGC CCAGCAACCT TGAAGGAGAT CAGCCACTGC AGGCAATCAT GGAGATGGGT 1890
 I D V P G Q V Q V Y E L Q P S N L E A D Q P L Q A I M E M G

 GCGCTGGAG CAGACAGGG CAAGAAATG CAGAAAGATCC CCACACAGAA ACCCAGGAGC CAGAAGCCAC AGCAGGAGG 1980
 A V A A D K G K N A G N A E D P H T E T Q Q P E A T A A A

 ACTTCAAAC CCAGCAGCAT GACAGACACC CCTGGTAACC CAGCAGCACC CCTAGCCTCTG CCCCTGTAAAAA ATCAGACTCG GAACGGATGT 2070
 T S N P S S M T D T P G N P A A P

 GTGCTTAGG GAATTAAAG TTGCTGCACT TTCAAGAAGCT TGAAGTCAGT TGGTTTTAT TAGCTGCTTG GTATGGAGA ACTTGGGG 2160
 AGGCAAAACA CTAATAAAAG GGCTAAAG GAAAATGATG CTTTCTCTT ATATTCTAC TCTGTCACAA TAAAGAAGTT GCTTGTGTT 2250
 TGAGAAGTT AACCCGTT CTGTTCTGC AGCCCTGCT ACTTGGCAC CCCACCAAC TGTAGCTGT GCTTGTGCACT TGTCTTTGT 2340
 AGCTCTGGAC TGGGGGATG CATTACCA TCACTAAAT ATGAAACATT TATCAGAAAT GTTGCCTATT TAATGAGATG 2430
 ATTTCTCA TCTCATATT AAAATACCTG ACTTTAGAGA GAGTAAATG TCCCAGGAGC CATAGGAATA TCTGATGTT GGATTGACTTT 2520
 ATGCTACAT TTTC

Fig. 15E

26/39

FIG. 16A

CGGTGGGAGC GGGGGGGAA GGGCTTCAAGG GCAGCGGATC CCATGTGGC
 CCTGAGGGGC TCGGGCTACG GCCCCAGTGA CGGTCCGGTCC TAGGGCCGCT 100
 ACTACGGGCC TGGGGTGGA GATGTGGG TACACCCACC TCCACCCCTTA 150
 TATCCTTC GCCCTGAACC TCCCCAGCCT CCCATTTCCT GGCGGGTGGC 200
 CGGGGGGGC CCGGGGAGA CCAACTGGCT GGAGGAAGGGC GGAGGGAGGG 250
 ATGGCTACTA TCCCTCGGGA GGCCTGGC CAGAGCCTGG TCGAGGGCGA 300
 GGAAGCCACC AGGAGCAGCC ACCCATATCCT AGCTACAATT CTAACATTG 350
 GAATTCTACT GCGAGATCTA GGGCTCCCTTA CCCAAGTACA TATCCTGTAA 400
 GACCAATT GCAAGGCCAG AGTTTGAATT CTTATACAAA TGGAGCCGAT 450
 GGTCCAACAT ACCCCCCAGG CCCTGGGCA AATACTGCCT CATACTCAGG 500
 GGCTTATTAT GCACCTGGTT ATACTCAGAC CAGTTACTCC ACAGAAGTTTC 550
 CAAGTACTTA CGGTTCATCT GGCAACAGCC CAACTCCAGT CTCTCGTTGG 600
 ATCTATCCCC AGCAGGACTG TCAGACTGAA GCACCCCCCT TTAGGGGCA 650
 GGTTCCAGGA TATCCGCCCTT CACAGAACCC TGGAAATGACC CTGCCCCATT 700
 ATCCTTATGG AGATGGTAAT CGTAGTGTTC CACAATCAGG ACCGACTGTA 750
 CGACCCACAAAG AAGATGGGTG GGCTTCTCCT GGTGCTTATG GAATGGGTGG 800
 CCGTTATCCC TGGCCCTCAT CAGCGGCCCTC AGCACCCACCC GGCAATCTCT 850
 ACATGACTGA AAGTACTTCA CCATGGCCTA GCAGTGGCTC TCCCCAGTCA 900
 CCCCCCTCAC CCCCAGTCCA GCAGCCCCAAG GATTCTTCAT ACCCCTATAG 950
 CCAATCAGAT CAAAGCATGA ACCGGCACAA CTTTCCCTTGC AGTGTCCATC 1000
 AGTACGAATC CTCGGGGACA GTGATCAATG AAGATTCAAG TCTTTGGAT 1050
 TCCCAAGTCC AGTATAAGTGC TGAGGCCCTCAG CTGTATGGTA ATGCCACACAG 1100
 TGACCATCCC AACAAATCAAG ATCAAAGTAG CAGTCTTCCT GAAGAAATGTG 1150
 TACCTTCAGA TGAAAGTACT CCTCCGGAGTA TTAAAAAAAT CATACTATGTG 1200
 CTGGAGAAGG TCCAGTATCT TGAACAAAGAA GTAGAAGAAAT TTGTAGGGAAA 1250
 AAAGACAGAC AAAGCATACT GGCTTCTGGA AGAAAATGCTA ACCAAGGAAC 1300

27/39

FIG. 16B

TTTGGAACT GGATTCAAGTT GAAACTGGGG GCCAGGACTC TGTACGGCAG 1350
GCCAGAAAAG AGGCTGTTG TAAGATTCAAG GCCATACTGG AAAAATTAGA 1400
AAAAAAAGGA TTATGAAAGG ATTTAGAACA AAGTGGAAAGC CTGTTACTAA 1450
CTTGACCCAA GAACACTTGA TTAGGTAAAT TACCCCTCTT TTGAAATGCC 1500
TGTGATGAC AAGAAGCAAT ACATTCAGC TTTTCCCTTG ATTTTATACT 1550
TGAAAAC TG GCAAAGGAAT GGAAGAATAT TTTAGTCATG AAGTTGGTTT 1600
CAGTTTCAGA CGAATGAATG TAATAGAAA CTATGGAGTT ACCAATATTG 1650
CCAAGTAGAC TCACTCCTTA AAAAATTAT GGATATCTAC AAGCTGCTTA 1700
TTACCAAGCAG GAGGGAAACA CACTTCACAC AACAGGGCTTA TCAGAAACCT 1750
ACCAGATGAA ACTGGATATA ATTTGAGACA AACAGGATGT GTTTTTTAA 1800
ACATCTGGAT ATCTTGTCAAC ATTTTGTAC ATTTGTGACTG CTTTCAACAT 1850
ATACTTCATG TGTAAATTATA GCTTAGACTT TAGCCTCTT GGACTTCTGT 1900
TTTGTTTGT TATTTGCAGT TTACAAATAT AGTATTATC TCTAAAAAA 1950
AAAAAAAGAA AAAAAGA 1966

28/39

FIG. 16C

MSAURRSGYGPSDGPSPYGRYYGGDVPVHPPPLYPLRPEPPQQPISWRVRGGGPAAETTWLGECCCCDGYYPSSGAWP
EPGRAGGSHQEQQPPSYNSNWNISTARSRAPYPSYNSPTVRAPELQGQSLNSYTNQAYGPTYPGPANTASYSGAYYAPGY
TQTSYSTEVPTSYRSGNSPTPVSRLWYPPQDQCTEAPLJRGQVPGYPPSCNPQGMTPHYGYDGNRSVPQSGPTVRPOE
DAWASPGAYGMGGRYPWPSSAPSAPPGNLYMTESTSPWPSSGSPQSQSPPPVQQPKDSSYPPYSQSDQSMNRHNFPCSWHQ
YESSGTVINEDSDLLDSQVQYSAEPQLYGNATSDHPNNDQSSSLPEECVPSDESTPPSIKKIIIHLEKVOYLEQEVEEF
VGKTDKAYWLLLEMLTKELLELDSVTGGQDSVRQARKEAVCKIQAILEKLEKKGL

29/39

CGTGGAGC GGGGGGAA GCGCTTCAAG GCAGGGATC CCTGAGGGCC CCTATGCGGC TCGGGCTACG GCCCCAGTGA CGGGCGTCC 90
 M S A L R R S G Y G P S D G P S

 TACGGGCT ACTACGGGCT TGGGGGCTGG GATGTCGGG TACACCCACC TCCACCCCTTA TATCCCTCTTC GCCCCAGGCT 160
 Y G R Y Y G P G G D V P V H P P P L Y P L R P E P P Q P

 CCCATTCCG GCGGGTGG CGGGGGGGC CGGGGGGAGA CCACCTGGCT GGGAGAAGGC GGAGGAGGC ATGGCTACTA TCCCTCGGGAA 270
 P I S W R V R G G P A E T T W L G E G G G D G Y Y P S G

 GGGCCCTGGC CAGAGCTGG TCGAGGGGA GGAAGCCACC AGGAGCAGCC ACCATATCCT AGCTACAATT CTAACATTG GAATTCTACT 360
 G A W P E P G R A G G S H Q E Q P Y P S Y N S N Y W N S T

 GCGAGATCTA GGGCTCTTA CCCAGTACA TATCCCTGTA GACCAAGATT GCAAGGGCAG AGTTTGAATT CTTATACAA TGGAGGGTAT 450
 A R S R A P Y P S T Y P V R P E L Q G Q S L N S Y T N G A Y

 GGTCCAAACAT ACCCCCCAGG CCTGGGCA AATACTGGCT CATACTCAGG GGCTTATTAT GCACCTGGTT ATACTCAGAC CAGTACTCC 540
 G P T Y P P G P G A N T A S Y S G A Y Y A P G Y T Q T S Y S

 ACAGAAGTTC CAAGTACTTA CGGTCATCT GGCAACAGCC CAACTCCAGT CTCTCGTTGG ATCTATCCCC AGCAGGACTG TCAGACTGAA 630
 T E V P S T Y R S S G N S P T P V S R W I Y P Q Q D C Q T E

 GCACCCCCCTC TTAGGGGCA GTTCCAGGA TATCCGCTT CACAGAACCC TGGAAATGACC CTGCCCCATT ATCCTTATGG AGATGGTAAT 720
 A P P L R G Q V P G Y P P S Q N P G M T L P H Y P Y G D G N

 CGTAGTGTTC CACATCAGG ACCGACTGTA CGACCAAG AAGATGCGTG GGCTTCTCT GGTCCTTATG GAATGGTGG CGGTATCCC 810
 R S V P Q S G P T V R P Q E D A W A S P G A Y G M G G R Y P

Fig. 16D

30/39

TGGCCCTCAT CAGCACCCCTC AGCACCAACCC GGCAATCTCT ACATGACTGA AAGTACTTCA CCATGGCCCTA GCAGTGGCTC TCCCCAGTCA
 W P S S A P P G N L Y M T E S T S P W P S S G S P Q S 900
 CCCCCCTTCAC CCCCCAGTCCAG GCAAGCCCAAG GATTCTTCAT ACCCCCTATAG CCAATCAGAT CAAAGCATGA ACCGGCACAA CTTTCTTGC
 P P S P P V Q P K D S S Y P Y S Q S D Q S M N R H N F P C 990
 AGTGTCCATC AGTACGAATC CTCGGGGACA GTGATCAATG AAGATTCAAGA TCTTTGGAT TCCCAAGTCC AGTATACTGC TGAGGCTCAG
 S V H Q Y E S S G T V I N E D S D L L D S Q V Q Y S A E P Q 1060
 CTGTATGTA ATGCCACCAAG TGACCATCCC AACAAATCAAG ATCAAAGTAG CAGTCTCT GAAGAATGTC TACCTTCAGA TGAAAGTACT
 L Y G N A T S D H P N N Q D Q S S S L P E E C V P S D E S T 1170
 CCTCCGAGTA TTAAAAAAAT CATACTATGTG CTGGAGAAGG TCCAGTATCT TGAACAAGAA GTAGAGAAAT TTGTAGGAA AAAGACAGAC
 P P S I K K I H V L E K V Q Y L E Q E V E F V G K K T D 1260
 AAAGCATACT GGCTTCTGGA AGAAATGCTA ACCAAGGAAAC TTTTGGAACT GGATTCAAGT GAAACTGGGG GCCAGGACTC TGTACGGCAG
 K A Y W L L E E M L T K E L L E L D S V E T G G Q D S V R Q 1350
 GCGAGAAAAAG AGGCTGTTG TAAGATTCAAG GGCATACTGG AAAAATTAGA AAAAAGGA TTATGAAAGG ATTAGAAACA AAGTGAAGC 1440
 A R K E A V C K I Q A I L E K L E K K G L
 CTGTTACTAA CTTGACCAA GAACACTTGA TTAGGTTAAT TACCTCTTT TTGAAATGCC TGTGATGAC AGAAAGCAAT ACATCCAGC
 TTTTCTTG ATTTTACT TGAAAACAG GCAAGGAAT TTAGTCATG AAGTTGTTT CAGTTTCAG AGAAATGAATG 1530
 TAATAGGAA CTATGGAGTT ACCAATATTG CCAGTAGAC TCACCTCTTA AAAAATTAT GGATATCTAC AGCTGCTTA TTACCAAGCAG
 GAGGGAAACA CACTCACAC AACAGGCTTA TCAGAAACCT ACCAGATGAA ACTGGATATA ATTTGAGACA ACAGGATGT GTTTTTTAA 1620
 ACATCTGGAT ATCTTGTAC ATTGTTGAC ATACTTCATG TGTAAATTATA GTTGTAGACTT TACCTCTT 1710
 GGACTTCTGT TTGTTTGT TATTGCACT AGTATTATTG TTACAAATAT TTACAAATTC TCTAAAAAA AAAA 1800
 1890
 1980

Fig. 16 E

31/39

FIG. 17A

CCCCCCCC CCCCCCCC CCGGAAGACG CCGGGAGCGG CTGCTGCAGC 50
CAGTAGGGC CCCCCACCG GCTGCCCGC TCAGACCTAG TCGGGAGGG 100
TGGGAGGCAT CGAGCTGGG GCGCAGCTCC GGTCGCGAC CCCGTAAGG 150
GCTGATCTTC CACCTGCCA CCTCAGCAC GGACGCCAA GACCGCATCC 200
AATTCAAGACT TCTTTGGTG CTTGTGAAAC TGAACACAAAC AAAAGTATGG 250
ATATGGAAA CCAACATCCT TCTATTAGTA GGCTTCAGTG GTCTGTCAGA 300
GAAGTAAAAA GTGTAGAACA GCAAGTTATC GGCTTCAGTG GTCTGTCAGA 350
TGACAAAGAAT TACAAGAAC TGGAGAGGAT TCTAACAAAA CAGCTTTTG 400
AAATAGACTC TGTAGATACT GAAGGAAAG GAGATATTCA GCAAGCTAGG 450
AAGGGGGCAG CACAGGAGAC AGAACGTCCT CTCAAAGAGT TGGAGCAGAA 500
TGCAAACAC CCACACCGGA TTGAAATACA GAACATTGTT GAGGAAGGCC 550
AGTOCCTCGT GAGAGAGAAA ATTGTGCCAT TTTATAATGG AGGCAACTGC 600
GTAACGTATG AGTTGAAGA AGGCATCCAA GATATCATTG TGAGGCTGAC 650
ACATGTTAAA ACTGGAGGAA AAATCTCCCT GCGGAAAGCA AGGTATCACA 700
CTTTAACAA AATCTGTGCG GTGCAAGAGA TAATCGAAGA CTGCATGAAA 750
AAGCAGCCTT CCCTGCCGCT TCCGAGGAT GCACATCCTT CCGTTGCCAA 800
AATCAACTTC GTGATGTGTG AGGTGAAC ACAATGAGA CCTGCAGGCA CTTATCCCTGT 850
CACTTCTGAT GGGTGTGAAC ACAATGAGA CCTGCAGGCA CTTATCCCTGT 900
GTGCTCTGG GGCTGATCGCTGACCTGGAT GCTCTAGATG TGTGGGGCG 950
GACAGAAATC AGAAAATTATC GGAGGGAGGGT AGTAGAAGAT ATCAACAAAT 1000
TATTGAAATA TCTGGATTTG GAAGAGGAAG CAGACACAAAC TAAAGCATT 1050
GACCTGAGAC AGAATCATTC CATTTTAAAAA ATAGAAAAGG TCCTCAAGAG 1100
AATGAGAGAA ATAAAAAATG AACTTCTCCA AGCACACAAAC CTTCTGTAAAT 1150
TGTACCTGAG CTCCAAACAA GAATTGCAGG GTTTAATTGG ACAGTTGGAT 1200
GAGGGTAAGTC TTGAAAAAAA CCCTGCATC CGGGAAAGCCA GGAGAAGAGC 1250
AGTGATCGAG GTGCAAAACTC TGATCACATA TATTGACTTG AAGGAGGCC 1300

32/39

FIG. 17B

TTGAGAAAAG AAAGCTGTTT GCTTGTGAGG AGCACCCATC CCATAAAGCC 1350
 GTCTGGAACG TCTTGGAAA CTTGCTGAG ATCCAGGGAG AAGTTCTTTTC 1400
 ATTTGATGGA AATCGAACCG ATAAGAACTA CATCCGGCTG GAAGAGCTGC 1450
 TCACCAAGCA GCTGCTAGCC CTGGATGCTG TTGATCCGCA GGGAGAAGAG 1500
 AAGTGTAGG CTGCCCAGGAA ACAAGCTGGTGA GGAGCTTGGCC AGAAATATTCT 1550
 CAGCTATCTC GACCTGAAAT CTGATGAATG GGAGTACTGA AATACCAAGAG 1600
 ATCTCACTTT TGATACTGTT TTGCACTTCA TATGTCCTC TATGTATAGA 1650
 GAGCTTCAG TTCAATTGATT TATACTGCA TATTTCAGTC TCAGTATTAA 1700
 TGATTGAAGC AAATTCTATT CAGTATCTGC TGCTTTTGAT GTTGCAGAC 1750
 AAATATCATT ACAGCACGTT AACTTTTCA TTGGGATCAT TATCTGTATG 1800
 ATGTGGTGTG GTTGTGTTGG TTTGTCCTT TTTTGCCTT TTTAATCAGA 1850
 AAACAAATA GAGGCAGCTT TTGTAGATT TAAATGGGTT GTGCAAGCAT 1900
 TAAAATGCAG GTCTTTCAAGA ATCTAGAACT AGGCATAACC TTACATAATA 1950
 CTAGGAAAAT TATGAGAAAG GGGAAATT TTGTTAAATA AGAGTAAGGT 2000
 TCAAAACACAA GCAGTACATG TTCTGTTCA TTATGCTGA TAGAAGGCTT 2050
 TTTTTCACT TATAAGGGCT GATTGGTCCT ACCCAGCTTA ACGGGGTGGG 2100
 GTTTTTGT TTGTTCAAGAC AGTCTGTTCT TTGTAAACA TTTTAGTTG 2150
 GAAAAACAGC ATCTGCATT TCCCCATCCT CTACGTTTA GAGGAGGAATC 2200
 TTGTTTTGT GTGCAACATA AGAAAATTAT GAAAACATAAT AGCCAAAAAA 2250
 CCTTTGAGAT TGCATTAAG AGAAGGGATA AAGGACCAAGC AATAATACCT 2300
 TGTAAAGTTGC TTTTGTGTTG AAAATCTGAG CTTATAGTTT CCCTTAGTGA 2350
 GTAAATTCAAT AAGGATGGGA ACATTTAAAT TAAGTTAATG GGCCCTTTAA 2400
 GAAAAAAAG GAAACACTCA TACCTGTAGT TGAGGATGA ATACTGGAGA 2450
 CGGGTTACCA ATGTCAGGGTT ATACTAAAC TAAATCAGAA AGTCTGAATG 2500
 TAGGCACATAA TGGTTCTCTT CTGGTGTCCA AGGCTGTAAA ATGGACAGGCC 2550
 TTGTCACACCA TCCCCGGTGC TGGTTTACAA CGTGAGGGTA GACCGCTGTCA 2600

FIG. 17C

GTAACCCAGA GGGACCAGGC CTTCCCTAGGT TTTCTAGGCC GTCAGGCTGTT
 ACCCACTCAC TTGTAATG TCATAACTAC ACCTGCTCCA GGACCAATCA 2650
 GTGAAACCTG CTCGGAAATTA AAGGCTTCTCT CTGGGTGCT GCTGAACAAAC
 TGAGCTCATG TCATGGGCAT GTGGTGGTT CTCTGTTGCC TGAAAGAGCC 2700
 ATTAAAGTCATG CTCGGTGGTG AAGCATTCTCT CTTCTAAAGG ATGTGTATT 2750
 CACATAATGC TTCTGAGGA TCCGGTACAA AATGATTTCC CAAAGTTCTG
 AAGTGCCTTG AGAACATGTTG GGTCCGGAGTG TTATAACAGA CTCCTCCCC 2800
 GGGTCAACCTT TTGCCTGTC ATCCCTGTTAG AGTACATCTT TGGAAATCCA
 GGGTAATATT CTCTTCAGA GATGCTCATT GTGTAACCT GTGTAGGGAG 2850
 ATAGTCACCTT TAAACAGCTC AAAGTAGCTA GCTAAAGGAG TAGCCTTAA
 TACCTAAAAG ATGACAGAAG CATAAGCCCTT AACAAATCTT CAGCTTGTCT 2900
 CTCAGTATTT CCCAATCATG AAAATCCCTT GCTATGTCCTT TCCTACTAGA
 ATATGTTCTAG AATCGCTGGA CGGTGGGTC AGAGGGCAGT CGGTATTAG 2950
 GCCGTGAGCT TCCCATACTA CTGCAGGTCC AACTCCTGGC AACCGGGGC
 TCAAGGCAGG TCATTGGAAAT CCACGTTTG GCCACAGTAG TTGTAGGATT 3000
 GCTTTCTGT ATCATAATTT TAGAATGCTC TTAAAATCTT GAGGAAGAGT
 TTTTATTTTT TATTTTTTT TGAGATGGAG TCTCTGTTGC CCAGGGCTGCA 3050
 GTGCAGTGGT GCCATCTCAG CTCACTGCAA CCTCCACCTC CCAGGGTCAA
 GCGGATTCTCC TGCCCTCAGCC ACCTGAGTAG CTGGGAGTAC AGGCATGTGG 3100
 CACCATGCCT GGCTAATTTT TGTTATTTA ATAGAGTTGA GATTTCACCA
 TGATGGTCAG GCTGGTCTCG AACTCCTGAC CTGGTGTGATCC GCCGGCCTCG
 GCCCCCCAA GTGCTGGGAT TAACGGGTGT GAGCCACGGC GCCCAGGCCA 3150
 GGAAGAGTT TAAATTAGA GCTCTGTTA ATTATACCACT TGGAAATCA
 TGGTTACGGCT TCAGGCATAT TCTTCCCCAG AGTACTACTT ACATTTAAA 3200
 TTTCAATTGG TAAAGTTAA GTGTCCATT CCCCTTAAAGA GTGTCCATTG
 TTCTTGGAAA GTAGACGGTT CAGTCATTCT CAGTCATTCT TTTCAAACAA GTGTGTGT 3250
 3300
 3350
 3400
 3450
 3500
 3550
 3600
 3650
 3700
 3750
 3800
 3850
 3900

34/39

FIG. 17D

ACCTTTGCC AAGCTGTGGG CATCGTGTGT GAGTACAGGG TGCTCAGCTC
TTCCACCGTC ATTGTAATT GTTCACATGG GTAATTGGTC ATGGAAATGA 3950
TCAGATTGAC CTTGATTGAC TGTCAAGGCAT GGCTTTGTTT CTAGTTCAA
TCTGTTCTCG TTCCCTGTAC CGGATTATTTC TACTCTGCA ATGAACCCCTG 4050
TTGACACCCGG ATTAGCTCT TGTCGGCCTT CGTGGGGAGCT GTTTGTGTT 4100
AATATGAGCT ACTGCATGTA ATTCTTAAAC TGGGCTTGTCA ACATTGTATT 4200
GTATTTTGT GATCTGTAAT GAAAAGAAATC TGTACTGCAG GTAAAACCTA 4250
CTCCCCAAA ATGTTGTGGCT TTGGGTCTGC ATTAACGGCT GTAGTCCATG 4300
TTCATGCC 4308

35 / 39

FIG. 17E

MDMGNQHPSI SRLQEIQKEV KSVEQQVIGF SGLSDDKNYK KLERLTKQL 50
FEIDSVDT EG KGDIQQARKR AAQETTERLKK ELEQNANHPH RIEIQNIFEE 100
AQSLVREKIV PFYNGGNCVT DEFEEGIIDI ILRLTHVKTG GKISLRKARY 150
HTLT KICAVQ EIIEDCMKKQ PSLPLSEDAAH PSVAKINFVM CEVNKARGVL 200
IALLMGVNN ETCRHLSCVL SGHLADLDAL DVCGRTEIRN YRREVVEDIN 250
KLLKYLDLEE EADTTKAFLR RQNHSLKIE KVLKRMREIK NELLQAQNPS 300
ELYLSSKTEL QGLIGQLDEV SLEKNPCIRE ARRRAVIEVQ TLTYIDLKE 350
ALEKRKLIFAC EEHPSHKAVW NVLGNLSEQ GEVLSFDGNR TDKNYIRLEE 400
LLTKQLLAID AVDPQGEEKC KAARKQAVRL AQNILSYLDL KSDEWEY 447

36/39

Fig. 17F

CCCCCCCCCC CCCCCCCCCC CNGAAGAG CCGGAGGG CTGGCTGCAGC CAGTAGGGC CCGCTTCAACGG GCTGCCCGG TCAAGACCTAG 90
 TGGGAGGGG TGGAGGCAT GCAGCTGGG GCCCAGCTCC GTGCGGCCAC CCCGTAAGG GCTGATCTTC CACCTGCCA CCTAGGCCAC 180
 GGGAGCCAA GACGGATCC ATTCAAGCT TCTTGTGGT CTTGTAAAC TGAAACAAAC AAAAGTATGG ATATGGAAA CCAACATCTT 270
 M D M G N Q H P

 TCTATTAGTA GCCTCAGGA ATCCAAAGA GAAGTAAAAA GTGTAGAACCA GCAAGTTATC GGCTTCAGTG GTCTGTAGA TGACAAGAAT 360
 S I S R L Q E I Q K E V K S V E Q Q V I G F S G L S D D K N
 TACAAGAAC TGGAGGAGT TCTAACAAA CAGCTTTTG AAATAGACTC TCTAGATACT GAAGGAAAG GAGATATTCA GCAAGCTAGG 450
 Y K K L E R I L T K Q L F E I D S V D T E G K G D I Q Q A R
 AAGGGGGAG CACAGGAGAC AGAACCGTCTT CTCAAAGAGT TGGAGGAGAA TGCAAAACAC CCACACCGGA TTGAAATACA GAACATTTTT 540
 K R A A Q E T E R L L K E L E Q N A N H P H R I E I Q N I F
 GAGGAAGCCC AGTCCCTGT GAGAGAGAAA ATTGTCCCAT TTTATAATGG AGGCAACTGG GTAACTGTAG AGTTTGAAAAG AGGCATCCAA 630
 E E A Q S L V R E K I V P F Y N G G N C V T D E F E E G I Q
 GATATCATTTC TGAGGCTGAC ACATGTAAAC ATGGAGGAA AAATCTCCTT GGGAAAGCA AGGTATCACCA CTTTAACCAA AATCTGTGCG 720
 D I I L R L T H V K T G G K I S L R K A R Y H T L T K I C A
 GTGCAAGAGA TAATGAAAGA CTGCATGAAA AAGCAGCCTT CCTCTGGCTT TTCCGAGGAT GCACATCCTT CGCTTGCCTT AATCAACTTC 810
 V Q E I I E D C M K K Q P S L P L S E D A H P S V A K I N F
 GTATGTGTG AGGTGAAACAA GGCCCCGAGGG GTCTCTGATT CACTTCTGAT GGGTGTGAAAC AACAATGAGA CCTGAGGCA CTTATCTGT 900
 V M C E V N K A R G V L I A L L M G V N N E T C R H L S C
 GTGCTCTGG GGCTGATGCC TGACCTGGAT GCTCTAGATG TGTCGGGGG GACAGAAATC AGAAATTATC GGAGGGAGGT AGTAAAGAT 990
 V L S G L I A D L D A L D V C G R T E I R N Y R R E V V E D
 ATCAACAAAT TATTGAAATA TCTGGATTG GAAGGAAAG AGAACACAACT TAAAGCATTG GACCTGAGAC AGAACATCTT CATTTTAAAA 1080
 I N K L L K Y L D L E E E A D T T K A F D L R Q N H S I L K

37/39

Fig. 17G

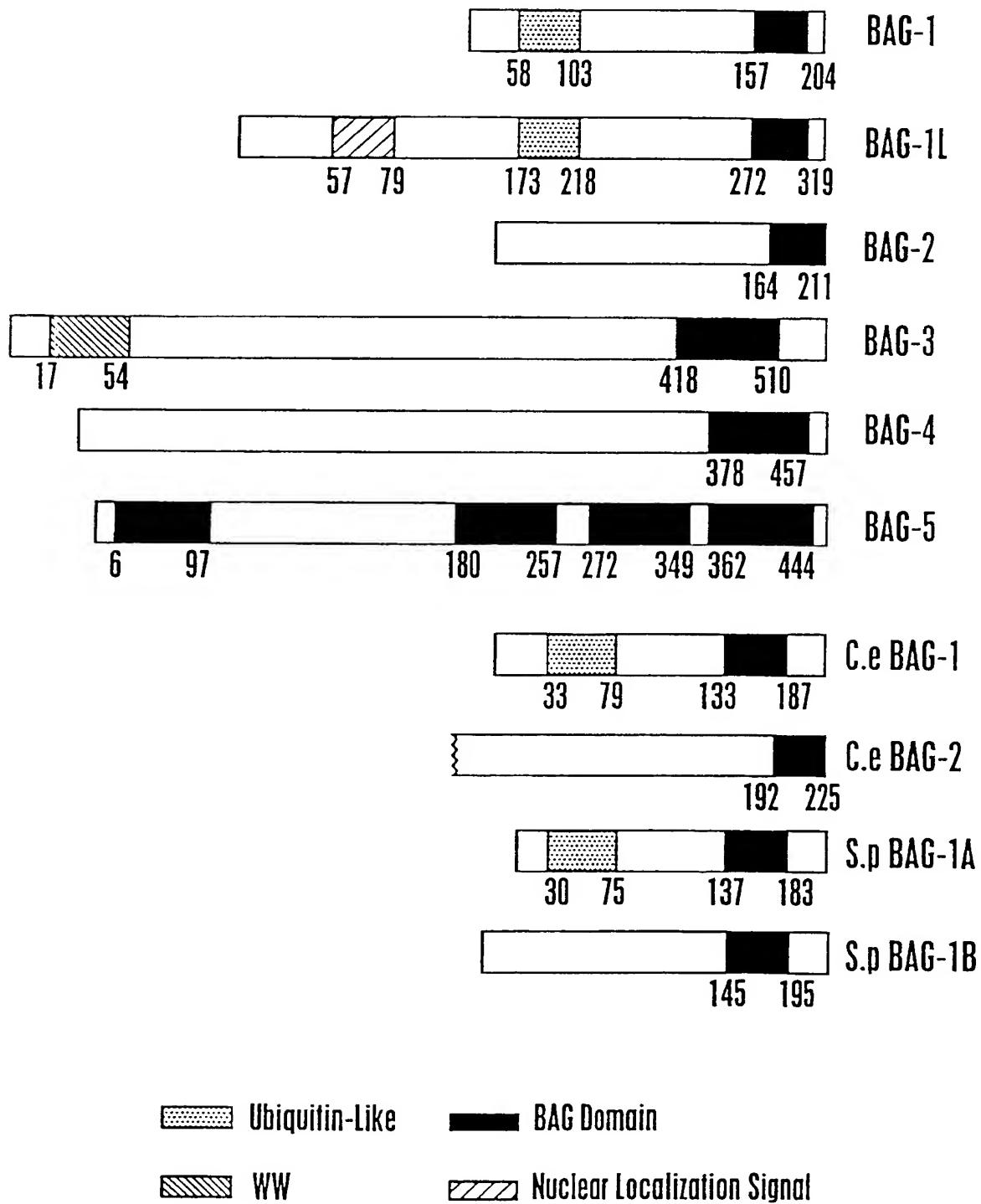
ATAGAAAAAGG TCTCTAACAGAG AATGAGAGAA ATAAAAAAATG AACTTCTCCA AGCACAAAAAC CCTTCTGAAT TGACCTGAG CTCACAAACA 1170
 I E K V L K R M R E I K N E L L Q A Q N P S E L Y L S S K T
 GAAATTGCAAGG GTTTAAATTGG ACAGTGGAT GAGGTAAGTC TTGAAAAAAA CCCCCTGCATC CGGGAAGCC GAGGAAAGAGC AGTGAATCGAG 1260
 E L Q G L I G Q L D E V S L E K N P C I R E A R R A V I E
 GTGCAAACTC TGATCACATA TATTGACTTG AAGGAGGCC TTGAGAAAAG AAAGCTGTTT GCTTGTGAGG AGCACCCATC CCATAAAGCC 1350
 V Q T L I T Y I D L K E A L E K R K L F A C E E H P S H K A
 GTCTGGAAAG TCCCTGGAAA CTTGCTGAG ATCCAGGGAG AAGTTCTTTC ATTGTATGGA AATCGAACCG ATAAGAACTA CATCGGGCTG 1440
 V W N V L G N L S E I Q G E V L S F D G N R T D K N Y I R L
 GAAATTGCAAGG GTTTAAATTGG ACAGTGGAT GAGGTAAGTC TTGAAAAAAA CCCCCTGCATC CGGGAAGCC GAGGAAAGAGC AGTGAATCGAG 1260
 E L Q G L I G Q L D E V S L E K N P C I R E A R R A V I E
 GTGCAAACTC TGATCACATA TATTGACTTG AAGGAGGCC TTGAGAAAAG AAAGCTGTTT GCTTGTGAGG AGCACCCATC CCATAAAGCC 1350
 V Q T L I T Y I D L K E A L E K R K L F A C E E H P S H K A
 GTCTGGAAAG TCCCTGGAAA CTTGCTGAG ATCCAGGGAG AAGTTCTTTC ATTGTATGGA AATCGAACCG ATAAGAACTA CATCGGGCTG 1440
 V W N V L G N L S E I Q G E V L S F D G N R T D K N Y I R L
 GAGAGCTGC TCACCAAGCA GCTGCTAGCC CTGGATGCTG TTGATCCGA GGGAGAAGAG AAGTGTAAAG CTGCCAGGAA ACAAGCTGTG 1530
 E E L L T K Q L L A L D A V D P Q G E E K C K A A R K Q A V
 AGGCTTGGC AGAATATTCT CAGCTATCTC GACCTGAAT CTGATGAATG GGAGTACTGA AATACCAAG AGTCACTTT TGATACTGTT 1620
 R L A Q N I L S Y L D L K S D E W E Y
 TTGCACTCA TATGTGCTTC TATGTATAGA GAGCTTCAG TTCTATTGATT TATACGTGCA TATTTCAGTC TCAGTATTAA TGATTGAAGC 1710
 AAATTCTATT CAGTATCTGC TGTCTTGTAT GTTGCAGAC AAATATCATT ACAGCACGTT AACTTTTCCA TTGGGATCAT TATCTGTATG 1800
 ATGTTGGTGTG GTTTGGTGG TTTGTCTT TTTTGGTT TTAAATCAGA AAACAAATA GAGGGAGCTT TTGTAGATT TAAATGGTT 1890
 GTGCAAGCAT TAAATGGAG GTCTTCTAGA ATCTAGAACT AGGCATAACCC TTACATATA CTAGGAAAAT TATGAGAAAG GGGAAATTTT 1980
 TGGTTAAATA AGAGTAAGGT TCAAACACAA GCACTGACATG TTATGCTGCA TAGAAGGCTT TTTTTCACT TATAAGGCT 2070

Fig. 17H

GATTGGCTT ACCCAGCTT ACGGGGGGGG GTTTTTTTGT TTGTTCAAGAC AGTCTGTTCT TTTGTAACAA TTTTAGTTG GAAAAACAGC
 ATCTGCATT TCCCCATCTT CTACGTTTA GAGAGGAATC TTGTTTTTGT GTGCAACATA AGAAAATTAT GAAACTAAT AGCCAAAAAA
 CCTTTGAGAT TGCATTAAAG AGAAGGGATA AAGGACCCAGC AATAATACCT TGTAAGTTGC TTTGTTTTGT AAAATCTGAG CTTATAGTT
 TCCCTTAGTGA GTAAATTCA AGGATGGGA ACATTAAAT TAAGTTAATG GGCCTTTAA AAAAAAAAG GAAACACTA TACCTGTAGT
 TGGAGGATGA ATACTGGAGA CGGGTTACCA ATGTCAGGTT ATACTAAAC TAAATCAGAA AGTCCTGAATG TAGCACATAA TGGTTCTCT
 CTGTTGTCAGCTCA AGGCTGTAAGA ATGGACAGCC TTGTCACACC TCCCCGGTGC TGTGTTACAA CGTGAGGGTA GACCGTGTCA GTAAACCCAGA
 GGGACCAAGGC CTTCTAGGT TTTCTAGGA CTCAAGCTGT AACCACTCAC TTAGTAATG TCATAACTAC ACCTGCTCCA GGACCAATCA
 GTGAAACCTG CTGGAAATT AAGGCTTCTT CTGGGTGCTTG TGAGCTCATG TGAGCAAC TGAGCAAC TGAGCTCATG TGAGCTCATG
 TGAAAGAGCC ATTAAGTCAGCTCA GTCTGTCGTT AAGGCTCTCT CTCTAAAGG ATGCTATT CCATAAATGC TTGCTGAGGA TCGGGTACAA
 AATGATTTC CAAAGTTCTG AAGTGCCTTG AGAACATGTG GGTCGGAGTG TTATAACAGA CTCTCTCCCC GGGTACCTT TTGCTGGTC
 ATCTGTTAG AGTACATCTT TGAATTCAGA GGGTAAATT CTCTTCAGA GATGCTATT GTGCTATT GTGTAACCTT GTGTAACCTT
 TAAACAGGTC AAAAGTAGCTA GCTAAAGGAG TAGCCCTAAAC TACCTAAAG ATGACAGAAG CATAGCCCTT ACAAATCTT CAGCTGCTT
 CTCACTATT CCCAATCTG AAATCCCTT GCTATGCTT TCCCTACTAGA AATGTTCTAG AATGCTGAA CGGTGGGTC AGAGGGCAGT
 CGGTATTAG GCGGTGAGCT TCCCCTACTA CTGCAAGTCC AACCTCTGGC AACGGGGGC TCAAGGCAAG TCATTGAAAT CCAGGTTTG
 GCCACAGTAG TTGTTAGATT GCTTTCTGT ATCATTAATT TGAATGCTC TTAAATCTT GAGGAAGAGT TTATTATTT TATTATTTT
 TGAGATGGAG TCTCTGTTGC CAGGGCTGCA GTGCACTGGT GCACTCTCA GCTACTGCAA CCTACTGCAA CCAGGTCCTC CGGATTCTCC
 TGCTCTAGCC ACCTGAGTAG CTGGGAGTAC AGGCATGTGG CACCATGCTT GGCTAAATT TGTTATTAA ATAGAGTTGA GATTCA
 TGATGGCTAG GCTGCTCTG ACTCCTGAC CTCGTGATCC GCGGGCTCG GCCCCAAA GTGCTGGGAT TAACGGGTG GAGCCACGGGC
 GCGGAGGCCA GGAAGAGTT TTAAATTAGA GCTCTGTTA ATTATACAC TGGAAATCA TGTTACGCT TCAGGCAAT TCTTCCCCAG
 AGTACTACTT ACATTAAAC TTTCAAAACAA GTGTTCTGT ACCTTTGCC AACCTGTGG CACCTGTTAAAGTTAA TGTCAGCATT
 CAGTCATTCT TTCAAAACAA GTGTTCTGT ACCTTTGCC AACCTGTGG CACCTGTTAAAGTTAA TGTCAGCATT
 ATTTGAATT GTTCACATGG GAAATTGGTC ATGAAATGA TCAAGATTGAC CTTGATTGAC TGTCAGGCT GTCAGGCT
 TCTGTTCTGG TTCCCTGTCAG CGGATTATC TACTCTGCA ATGAAACCTG TTGACACGGG ATTAGCTCT TGTGCCCCCTT
 TGTGTTGCTT AATATGAGCT ACTGCAATGA ATTCTAAAC TGGGCTTGT ACATTGATT GTATTTGT GATCTGTAAT GAAAAGAATC
 TGTACTGCAA GTAAACACTA CTCCCCAAAATGTTGGCT TTGTTGGCT ATTAAACGCT GTAGTCCATG TTGATGCC

39/39

Fig. 18



SEQUENCE LISTING

<110> Reed, John C.

Takayama, Shinichi

The Burnham Institute

<120> Novel BAG Proteins and Nucleic Acid Molecules Encoding
Them

<130> FP-LJ 3646

<140>

<141>

<150> 09/150,489

<151> 1998-09-09

<160> 24

<170> PatentIn Ver. 2.0

<210> 1

<211> 1291

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (46)..(1080)

<400> 1

acgcccgcgt cagttccat cgctgggcgg tcaacaagtg cgggc ctg gct cag cgc 57

Leu Ala Gln Arg

1

ggg ggg gcg cgg aga ccg cga ggc gac cgg gag cgg ctg ggt tcc cgg 105
Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg Leu Gly Ser Arg

5

10

15

20

ctg cgc gcc ctt cgg cca ggc cgg gag ccg cgc cag tcg gag ccc ccc 153
Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln Ser Glu Pro Pro

25

30

35

gcc cag cgt ggt ccg cct ccc tct cgg cgt cca cct gcc cgg agt act 201
Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro Ala Arg Ser Thr

40

45

50

gcc agc ggg cat gac cga ccc acc agg ggc gcc gcc ggc gct cgc 249

Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala Ala Gly Ala Arg
 55 60 65

agg ccg cgg atg aag aag aaa acc cgg cgc cgc tcg acc acc cgg agc gag 297
 Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser Thr Arg Ser Glu
 70 75 80

gag ttg acc cgg agc gag gag ttg acc ctg agt gag gaa gca acc tgg 345
 Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu Glu Ala Thr Trp
 85 90 95 100

agt gaa gag gca acc cag agt gag gag gca acc cag ggc gaa gag atg 393
 Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln Gly Glu Glu Met
 105 110 115

aat cgg agc cag gag gtg acc cgg gac gag gag tcg acc cgg agc gag 441
 Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser Thr Arg Ser Glu
 120 125 130

gag gtg acc agg gag gaa atg gca gct ggg ctc acc gtg act gtc 489
 Glu Val Thr Arg Glu Glu Met Ala Ala Ala Gly Leu Thr Val Thr Val
 135 140 145

acc cac agc aat gag aag cac gac ctt cat gtt acc tcc cag cag ggc 537
 Thr His Ser Asn Glu Lys His Asp Leu His Val Thr Ser Gln Gln Gly
 150 155 160

agc agt gaa cca gtt gtc caa gac ctg gcc cag gtt gtt gaa gag gtc 585
 Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val Val Glu Glu Val
 165 170 175 180

ata ggg gtt cca cag tct ttt cag aaa ctc ata ttt aag gga aaa tct 633
 Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe Lys Gly Lys Ser
 185 190 195

ctg aag gaa atg gaa aca ccg ttg tca gca ctt gga ata caa gat ggt 681
 Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly Ile Gln Asp Gly
 200 205 210

tgc cgg gtc atg tta att ggg aaa aag aac agt cca cag gaa gag gtt 729
 Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro Gln Glu Glu Val
 215 220 225

gaa cta aag aag ttg aaa cat ttg gag aag tct gtg gag aag ata gct 777
 Glu Leu Lys Lys Leu Lys His Leu Glu Lys Ser Val Glu Lys Ile Ala
 230 235 240

gac cag ctg gaa gag ttg aat aaa gag ctt act gga atc cag cag ggt 825

Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly Ile Gln Gln Gly
 245 250 255 260
 ttt ctg ccc aag gat ttg caa gct gaa gct ctc tgc aaa ctt gat agg 873
 Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys Lys Leu Asp Arg
 265 270 275
 aga gta aaa gcc aca ata gag cag ttt atg aag atc ttg gag gag att 921
 Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile Leu Glu Glu Ile
 280 285 290
 gac aca ctg atc ctg cca gaa aat ttc aaa gac agt aga ttg aaa agg 969
 Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser Arg Leu Lys Arg
 295 300 305
 aaa ggc ttg gta aaa aag gtt cag gca ttc cta gcc gag tgt gac aca 1017
 Lys Gly Leu Val Lys Val Gln Ala Phe Leu Ala Glu Cys Asp Thr
 310 315 320
 gtg gag cag aac atc tgc cag gag act gag cgg ctg cag tct aca aac 1065
 Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu Gln Ser Thr Asn
 325 330 335 340
 ttt gcc ctg gcc gag tgaggtgttag cagaaaaagg ctgtgctgcc ctgaagaatg 1120
 Phe Ala Leu Ala Glu
 345
 gcgccaccag ctctgccgtc tctggatcg aatttacctg atttcttcag ggctgctggg 1180
 ggcaactggc catttgccaa ttttcctact ctcacactgg ttctcaatga aaaatagtgt 1240
 ctttgatt tgagtaaagc tcctattctg ttttcacaa aaaaaaaaaa a 1291

<210> 2
 <211> 345
 <212> PRT
 <213> Homo sapiens

<400> 2
 Leu Ala Gln Arg Gly Gly Ala Arg Arg Pro Arg Gly Asp Arg Glu Arg
 1 5 10 15

Leu Gly Ser Arg Leu Arg Ala Leu Arg Pro Gly Arg Glu Pro Arg Gln
 20 25 30

Ser Glu Pro Pro Ala Gln Arg Gly Pro Pro Pro Ser Arg Arg Pro Pro
 35 40 45

Ala Arg Ser Thr Ala Ser Gly His Asp Arg Pro Thr Arg Gly Ala Ala
50 55 60

Ala Gly Ala Arg Arg Pro Arg Met Lys Lys Lys Thr Arg Arg Arg Ser
65 70 75 80

Thr Arg Ser Glu Glu Leu Thr Arg Ser Glu Glu Leu Thr Leu Ser Glu
85 90 95

Glu Ala Thr Trp Ser Glu Glu Ala Thr Gln Ser Glu Glu Ala Thr Gln
100 105 110

Gly Glu Glu Met Asn Arg Ser Gln Glu Val Thr Arg Asp Glu Glu Ser
115 120 125

Thr Arg Ser Glu Glu Val Thr Arg Glu Glu Met Ala Ala Ala Gly Leu
130 135 140

Thr Val Thr Val Thr His Ser Asn Glu Lys His Asp Leu His Val Thr
145 150 155 160

Ser Gln Gln Gly Ser Ser Glu Pro Val Val Gln Asp Leu Ala Gln Val
165 170 175

Val Glu Glu Val Ile Gly Val Pro Gln Ser Phe Gln Lys Leu Ile Phe
180 185 190

Lys Gly Lys Ser Leu Lys Glu Met Glu Thr Pro Leu Ser Ala Leu Gly
195 200 205

Ile Gln Asp Gly Cys Arg Val Met Leu Ile Gly Lys Lys Asn Ser Pro
210 215 220

Gln Glu Glu Val Glu Leu Lys Leu Lys His Leu Glu Lys Ser Val
225 230 235 240

Glu Lys Ile Ala Asp Gln Leu Glu Glu Leu Asn Lys Glu Leu Thr Gly
245 250 255

Ile Gln Gln Gly Phe Leu Pro Lys Asp Leu Gln Ala Glu Ala Leu Cys
260 265 270

Lys Leu Asp Arg Arg Val Lys Ala Thr Ile Glu Gln Phe Met Lys Ile
275 280 285

Leu Glu Glu Ile Asp Thr Leu Ile Leu Pro Glu Asn Phe Lys Asp Ser
290 295 300

Arg Leu Lys Arg Lys Gly Leu Val Lys Lys Val Gln Ala Phe Leu Ala
 305 310 315 320

Glu Cys Asp Thr Val Glu Gln Asn Ile Cys Gln Glu Thr Glu Arg Leu
 325 330 335

Gln Ser Thr Asn Phe Ala Leu Ala Glu
 340 345

<210> 3

<211> 1179

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (160)...(792)

<400> 3

gcagccgcgg tgcgcgaag tcctccggg ttgccccgc ggcgtcagag ggagggcggg 60

cgcgcgttg gtgacggcga ccctgcagcc caaggagcgc tccactcgct gcccggag 120

ggccgggtgac ctcttgcta cccgcgtcg gaggcttag atg gct cag gcg aag 174
 Met Ala Gln Ala Lys
 1 5

atc aac gct aaa gcc aac gag ggg cgc ttc tgc cgc tcc tcc atg 222
 Ile Asn Ala Lys Ala Asn Glu Gly Arg Phe Cys Arg Ser Ser Met
 10 15 20

gct gac cgc tcc agc cgc ctg ctg gag agc ctg gac cag ctg gag ctc 270
 Ala Asp Arg Ser Ser Arg Leu Leu Glu Ser Leu Asp Gln Leu Glu Leu
 25 30 35

agg gtt gaa gct ttg aga gaa gca gca act gct gtt gag caa gag aaa 318
 Arg Val Glu Ala Leu Arg Glu Ala Ala Thr Ala Val Glu Gln Glu Lys
 40 45 50

gaa atc ctt ctg gaa atg atc cac agt atc caa aat agc cag gac atg 366
 Glu Ile Leu Leu Glu Met Ile His Ser Ile Gln Asn Ser Gln Asp Met
 55 60 65

agg cag atc agt gac gga gaa aga gaa gaa tta aat ctg act gca aac 414
 Arg Gln Ile Ser Asp Gly Glu Arg Glu Glu Leu Asn Leu Thr Ala Asn
 70 75 80 85

cgt ttg atg gga aga act ctc acc gtt gaa gtg tca gta gaa aca att 462
 Arg Leu Met Gly Arg Thr Leu Thr Val Glu Val Ser Val Glu Thr Ile
 90 95 100

aga aac ccc cag cag caa gaa tcc cta aag cat gcc aca agg att att 510
 Arg Asn Pro Gln Gln Gln Glu Ser Leu Lys His Ala Thr Arg Ile Ile
 105 110 115

gat gag gat gtc aat aag ttt ctg gat gat ttg gga aat gcc aag agt 558
 Asp Glu Val Val Asn Lys Phe Leu Asp Asp Leu Gly Asn Ala Lys Ser
 120 125 130

cat tta atg tcg ctc tac agt gca tgt tca tct gag gtg cca cat ggg 606
 His Leu Met Ser Leu Tyr Ser Ala Cys Ser Ser Glu Val Pro His Gly
 135 140 145

cca gtt gat cag aag ttt caa tcc ata gta att ggc tgt gct ctt gaa 654
 Pro Val Asp Gln Lys Phe Gln Ser Ile Val Ile Gly Cys Ala Leu Glu
 150 155 160 165

gat cag aag aaa att aag aga aga tta gag act ctg ctt aga aat att 702
 Asp Gln Lys Lys Ile Lys Arg Arg Leu Glu Thr Leu Leu Arg Asn Ile
 170 175 180

gaa aac tct gac aag gcc atc aag cta tta gag cat tct aaa gga gct 750
 Glu Asn Ser Asp Lys Ala Ile Lys Leu Leu Glu His Ser Lys Gly Ala
 185 190 195

ggt tcc aaa act ctg caa caa aat gct gaa agc aga ttc aat 792
 Gly Ser Lys Thr Leu Gln Gln Asn Ala Glu Ser Arg Phe Asn
 200 205 210

tagtcttcaa acctaagagc attcacacaa tacacaaggt gtaaaaatga taaaatacta 852

ttttaattga taactagttc ttgttaggt ataaccactt agttgacact gatagttgtt 912

tcagatgagg aaaatattcc atcaagtatc ttcagtttg tgaataacaa aactagcaat 972

attttaatta tctatctaga gatTTTTtag attgaattct tgtcttgac taggatctag 1032

catatttcaatattctgtgg atgaatacat agtttgtggg gaaaacaaac gttcagctag 1092

gggcaaaaag catgactgct tttcctgtc tggcatggaa tcacgcagtc accttggca 1152

tttagtttac tagaaattct ttactgg 1179

<210> 4

<211> 211

<212> PRT

<213> Homo sapiens

<400> 4

Met	Ala	Gln	Ala	Lys	Ile	Asn	Ala	Lys	Ala	Asn	Glu	Gly	Arg	Phe	Cys
1				5					10					15	

Arg	Ser	Ser	Ser	Met	Ala	Asp	Arg	Ser	Ser	Arg	Leu	Leu	Glu	Ser	Leu
				20					25				30		

Asp	Gln	Leu	Glu	Leu	Arg	Val	Glu	Ala	Leu	Arg	Glu	Ala	Ala	Thr	Ala
					35			40				45			

Val	Glu	Gln	Glu	Lys	Glu	Ile	Leu	Leu	Glu	Met	Ile	His	Ser	Ile	Gln
					50			55			60				

Asn	Ser	Gln	Asp	Met	Arg	Gln	Ile	Ser	Asp	Gly	Glu	Arg	Glu	Glu	Leu
				65			70			75			80		

Asn	Leu	Thr	Ala	Asn	Arg	Leu	Met	Gly	Arg	Thr	Leu	Thr	Val	Glu	Val
					85				90				95		

Ser	Val	Glu	Thr	Ile	Arg	Asn	Pro	Gln	Gln	Glu	Ser	Leu	Lys	His
				100				105			110			

Ala	Thr	Arg	Ile	Ile	Asp	Glu	Val	Val	Asn	Lys	Phe	Leu	Asp	Asp	Leu
				115			120				125				

Gly	Asn	Ala	Lys	Ser	His	Leu	Met	Ser	Leu	Tyr	Ser	Ala	Cys	Ser	Ser
				130			135			140					

Glu	Val	Pro	His	Gly	Pro	Val	Asp	Gln	Lys	Phe	Gln	Ser	Ile	Val	Ile
				145			150			155			160		

Gly	Cys	Ala	Leu	Glu	Asp	Gln	Lys	Lys	Ile	Lys	Arg	Arg	Leu	Glu	Thr
					165				170			175			

Leu	Leu	Arg	Asn	Ile	Glu	Asn	Ser	Asp	Lys	Ala	Ile	Lys	Leu	Leu	Glu
				180				185			190				

His	Ser	Lys	Gly	Ala	Gly	Ser	Lys	Thr	Leu	Gln	Gln	Asn	Ala	Glu	Ser
					195			200			205				

Arg	Phe	Asn
-----	-----	-----

210

<210> 5
 <211> 2528
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(2031)

<400> 5
 gcg gag ctc cgc atc caa ccc cgg gcc gcg gcc aac ttc tct gga ctg 48
 Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Asn Phe Ser Gly Leu
 1 5 10 15

gac cag aag ttt cta gcc ggc cag ttg cta cct ccc ttt atc tcc tcc 96
 Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser
 20 25 30

ttc ccc tct ggc agc gag gag gct att tcc aga cac ttc cac ccc tct 144
 Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser
 35 40 45

ctg gcc acg tca ccc ccg cct tta att cat aaa ggt gcc cgg cgc cgg 192
 Leu Ala Thr Ser Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg
 50 55 60

ctt ccc gga cac gtc ggc ggc gga gag ggg ccc acg gcg gcg gcc cgg 240
 Leu Pro Gly His Val Gly Gly Glu Gly Pro Thr Ala Ala Ala Arg
 65 70 75 80

cca gag act cgg cgc ccg gag cca gcg ccc cgc acc cgc gcc cca gcg 288
 Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala
 85 90 95

ggc aga ccc caa ccc agc atg agc gcc gcc acc cac tcg ccc atg atg 336
 Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met
 100 105 110

cag gtg gcg tcc ggc aac ggt gac cgc gac cct ttg ccc ccc gga tgg 384
 Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp
 115 120 125

gag atc aag atc gac ccg cag acc ggc tgg ccc ttc ttc gtg gac cac 432
 Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His
 130 135 140

aac agc cgc acc act acg tgg aac gac ccg cgc gtg ccc tct gag ggc 480

Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly			
145	150	155	160
ccc aag gag act cca tcc tct gcc aat ggc cct tcc cgg gag ggc tct 528			
Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser			
165	170	175	
agg ctg ccg cct gct agg gaa ggc cac cct gtg tac ccc cag ctc cga 576			
Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg			
180	185	190	
cca ggc tac att ccc att cct gtg ctc cat gaa ggc gct gag aac cgg 624			
Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg			
195	200	205	
cag gtg cac cct ttc cat gtc tat ccc cag cct ggg atg cag cga ttc 672			
Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe			
210	215	220	
cga act gag gcg gca gca gcg gct cct cag agg tcc cag tca cct ctg 720			
Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu			
225	230	235	240
cgg ggc atg cca gaa acc act cag cca gat aaa cag tgt gga cag gtg 768			
Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val			
245	250	255	
gca gcg gcg gca gcc cag ccc cca gcc tcc cac gga cct gag cgg 816			
Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg			
260	265	270	
tcc cag tct cca gct gcc tct gac tgc tca tcc tca tcc tcc tcg gcc 864			
Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ala			
275	280	285	
agc ctg cct tcc tcc ggc agg agc agc ctg ggc agt cac cag ctc ccg 912			
Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro			
290	295	300	
cgg ggg tac atc tcc att ccg gtg ata cac gag cag aac gtt acc cgg 960			
Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg			
305	310	315	320
cca gca gcc cag ccc tcc ttc cac aaa gcc cag aag acg cac tac cca 1008			
Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro			
325	330	335	
gcg cag agg ggt gag tac cag acc cac cag cct gtg tac cac aag atc 1056			

Ala Gln Arg Gly Glu Tyr Gin Thr His Gln Pro Val Tyr His Lys Ile
 340 345 350

cag ggg gat gac tgg gag ccc cgg ccc ctg cgg gcg gca tcc ccg ttc 1104
 Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe
 355 360 365

agg tca tct gtc cag ggt gca tcg agc cgg gag ggc tca cca gcc agg 1152
 Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg
 370 375 380

agc agc acg cca ctc cac tcc ccc tcg ccc atc cgt gtg cac acc gtg 1200
 Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val
 385 390 395 400

gtc gac agg cct cag cag ccc atg acc cat cga gaa act gca cct gtt 1248
 Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val
 405 410 415

tcc cag cct gaa aac aaa cca gaa agt aag cca ggc cca gtt gga cca 1296
 Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro
 420 425 430

gaa ctc cct cct gga cac atc cca att caa gtg atc cgc aaa gag gtg 1344
 Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val
 435 440 445

gat tct aaa cct gtt tcc cag aag ccc cca cct ccc tct gag aag gta 1392
 Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Ser Glu Lys Val
 450 455 460

gag gtg aaa gtt ccc cct gct cca gtt cct tgt cct ccc agc cct 1440
 Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Ser Pro
 465 470 475 480

ggc cct tct gct gtc ccc tct tcc ccc aag agt gtg gct aca gaa gag 1488
 Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu
 485 490 495

agg gca gcc ccc agc act gcc cct gca gaa gct aca cct cca aaa cca 1536
 Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro
 500 505 510

gga gaa gcc gag gct ccc cca aaa cat cca gga gtg ctg aaa gtg gaa 1584
 Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu
 515 520 525

gcc atc ctg gag aag gtg cag ggg ctg gag cag gct gta gac aac ttt 1632

Aia Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe
 530 535 540 1680
 gaa ggc aag aag act gac aaa aag tac ctg atg atc gaa gag tat ttg
 Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu
 545 550 555 560
 acc aaa gag ctg ctg gcc ctg gat tca gtg gac ccc gag gga cga gcc 1728
 Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala
 565 570 575
 gat gtg cgt cag gcc agg aga gac ggt gtc agg aag gtt cag acc atc 1776
 Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile
 580 585 590
 ttg gaa aaa ctt gaa cag aaa gcc att gat gtc cca ggt caa gtc cag 1824
 Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln
 595 600 605
 gtc tat gaa ctc cag ccc agc aac ctt gaa gca gat cag cca ctg cag 1872
 Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln
 610 615 620
 gca atc atg gag atg ggt gcc gtg gca gca gac aag ggc aag aaa aat 1920
 Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn
 625 630 635 640
 gct gga aat gca gaa gat ccc cac aca gaa acc cag cag cca gaa gcc 1968
 Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala
 645 650 655
 aca gca gca gcg act tca aac ccc agc agc atg aca gac acc cct ggt 2016
 Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly
 660 665 670
 aac cca gca gca ccg tagcctctgc cctgtaaaag tcagactcgg aaccgatgtg 2071
 Asn Pro Ala Ala Pro
 675
 tgcttaggg attttagttt catgcatttc agagacttta ggtcagttgg tttttagttt 2131
 ctgcctggta tgcgtactt gggtagggca aacactataa agggctaaaa gggaaaatga 2191
 tgctttctt caatattctt actcttgtac aatataangaa gttgcttgg ttttgagaag 2251
 tttaaccccg ttgcttggc tgcagccctg tcnacttggg caccccccacc acctgttagc 2311
 tgtggttggc cactgtctt tgcgtctg gactggaggc gtagatgggg agtcaattac 2371

ccatcacata aatatgaaac atttatcaga aatgttgcca ttttaatgag atgattttct 2431
tcatctcata attaaaatac ctgactttag agagagtaaa atgtgccagg agccatagga 2491
atatctgtat gttggatgac tttaatgcta catttth 2528

<210> 6
<211> 677
<212> PRT
<213> Homo sapiens

<400> 6
Ala Glu Leu Arg Ile Gln Pro Arg Ala Ala Asn Phe Ser Gly Leu
1 5 10 15

Asp Gln Lys Phe Leu Ala Gly Gln Leu Leu Pro Pro Phe Ile Ser Ser
20 25 30

Phe Pro Ser Gly Ser Glu Glu Ala Ile Ser Arg His Phe His Pro Ser
35 40 45

Leu Ala Thr Ser Pro Pro Pro Leu Ile His Lys Gly Ala Arg Arg Arg
50 55 60

Leu Pro Gly His Val Gly Gly Glu Gly Pro Thr Ala Ala Ala Arg
65 70 75 80

Pro Glu Thr Arg Arg Pro Glu Pro Ala Pro Arg Thr Arg Ala Pro Ala
85 90 95

Gly Arg Pro Gln Pro Ser Met Ser Ala Ala Thr His Ser Pro Met Met
100 105 110

Gln Val Ala Ser Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp
115 120 125

Glu Ile Lys Ile Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His
130 135 140

Asn Ser Arg Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly
145 150 155 160

Pro Lys Glu Thr Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser
165 170 175

Arg Leu Pro Pro Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg

180	185	190
Pro Gly Tyr Ile Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg		
195	200	205
Gln Val His Pro Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe		
210	215	220
Arg Thr Glu Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu		
225	230	235
Arg Gly Met Pro Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val		
245	250	255
Ala Ala Ala Ala Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg		
260	265	270
Ser Gln Ser Pro Ala Ala Ser Asp Cys Ser Ser Ser Ser Ser Ala		
275	280	285
Ser Leu Pro Ser Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro		
290	295	300
Arg Gly Tyr Ile Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg		
305	310	315
320		
Pro Ala Ala Gln Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro		
325	330	335
Ala Gln Arg Gly Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile		
340	345	350
Gln Gly Asp Asp Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe		
355	360	365
Arg Ser Ser Val Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg		
370	375	380
Ser Ser Thr Pro Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val		
385	390	395
400		
Val Asp Arg Pro Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val		
405	410	415
Ser Gln Pro Glu Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro		
420	425	430
Glu Leu Pro Pro Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val		

435 440 445

Asp Ser Lys Pro Val Ser Gln Lys Pro Pro Pro Pro Ser Glu Lys Val
450 455 460

Glu Val Lys Val Pro Pro Ala Pro Val Pro Cys Pro Pro Pro Ser Pro
465 470 475 480

Gly Pro Ser Ala Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu
485 490 495

Arg Ala Ala Pro Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro
500 505 510

Gly Glu Ala Glu Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu
515 520 525

Ala Ile Leu Glu Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe
530 535 540

Glu Gly Lys Lys Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu
545 550 555 560

Thr Lys Glu Leu Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala
565 570 575

Asp Val Arg Gln Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile
580 585 590

Leu Glu Lys Leu Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln
595 600 605

Val Tyr Glu Leu Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln
610 615 620

Ala Ile Met Glu Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn
625 630 635 640

Ala Gly Asn Ala Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala
645 650 655

Thr Ala Ala Ala Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly
660 665 670

Asn Pro Ala Ala Pro
675

<210> 7
<211> 1010
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (323)..(1009)

<400> 7
acgatatacct gtaagaccaa gaattgcaag gccagagttt gaattcttat acaaatggag 60
cgtatggtcc aacataccccc ccagggcctg gggcaaatac tgcctcatac tcaggggctt 120
attatgcacc tggttatact cagaccagtt actccacacaga agttccaagt acttaccgtt 180
catctggcaa cagcccaact ccagtctctc gttggatcta tccccagcag gactgtcaag 240
actgaagcac cccctcttaa gggcaggtt ccaggatatc cgccttcaca gaaccctgga 300
atgaccctgc cccattatcc tt atg gag atg gta atc gta gtg ttc cac aat 352
Met Glu Met Val Ile Val Val Phe His Asn
1 5 10
15 20 25
cac ggc cga ctg tac gac cac aag aaa gat gcg tgg gct tct cct ggt 400
His Gly Arg Leu Tyr Asp His Lys Lys Asp Ala Trp Ala Ser Pro Gly
30 35 40
gct tat gga atg ggt ggc cgt tat ccc tgg cct tca tca gcg ccc tca 448
Ala Tyr Gly Met Gly Gly Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser
45 50 55
gca cca ccc ggc aat ctc tac atg act gaa agt act tca cca tgg cct 496
Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro
60 65 70
agc agt ggc tct ccc cag tca ccc cct tca ccc cca gtc cag cag ccc 544
Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val Gln Gln Pro
75 80 85 90
aag gat tct tca tac ccc tat agc caa tca gat caa agc atg aac cgg 592
Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg
95 100 105
cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tgc ggg aca gtg 640
His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser Gly Thr Val

aac aat gat gat tca gat ctt ttg gat tcc caa gtc cag tat agt gct 688
 Asn Asn Asp Asp Ser Asp Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala
 110 115 120

gag cct cag ctg tat ggt aat gcc acc agt gac cat ccc aac aat caa 736
 Glu Pro Gln Leu Tyr Gly Asn Ala Thr Ser Asp His Pro Asn Asn Gln
 125 130 135

gat caa agt agc agt ctt cct gaa gaa tgt gta cct tca gat gaa agt 784
 Asp Gln Ser Ser Leu Pro Glu Glu Cys Val Pro Ser Asp Glu Ser
 140 145 150

act cct ccg agt att aaa aaa atc ata cat gtg ctg gag aag gtc cag 832
 Thr Pro Pro Ser Ile Lys Lys Ile Ile His Val Leu Glu Lys Val Gln
 155 160 165 170

tat ctt gaa caa gaa gta gaa gaa ttt gta gga aaa aag aca gac aaa 880
 Tyr Leu Glu Gln Glu Val Glu Glu Phe Val Gly Lys Lys Thr Asp Lys
 175 180 185

gca tac tgg ctt ctg gaa gaa atg cta acc aag gaa ctt ttg gaa ctg 928
 Ala Tyr Trp Leu Leu Glu Glu Met Leu Thr Lys Glu Leu Leu Glu Leu
 190 195 200

gat tca gtt gaa act ggg ggc cag gac tct gta cgg cag gcc aga aaa 976
 Asp Ser Val Glu Thr Gly Gly Gln Asp Ser Val Arg Gln Ala Arg Lys
 205 210 215

gag gct gtt tgt aag att cag gcc ata ttg gaa a 1010
 Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu
 220 225

<210> 8

<211> 229

<212> PRT

<213> Homo sapiens

<400> 8

Met Glu Met Val Ile Val Val Phe His Asn His Gly Arg Leu Tyr Asp
 1 5 10 15

His Lys Lys Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly
 20 25 30

Arg Tyr Pro Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu
 35 40 45

Tyr Met Thr Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln
50 55 60

Ser Pro Pro Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro
65 70 75 80

Tyr Ser Gln Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser
85 90 95

Val His Gln Tyr Glu Ser Ser Gly Thr Val Asn Asn Asp Asp Ser Asp
100 105 110

Leu Leu Asp Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly
115 120 125

Asn Ala Thr Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu
130 135 140

Pro Glu Glu Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys
145 150 155 160

Lys Ile Ile His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val
165 170 175

Glu Glu Phe Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu
180 185 190

Glu Met Leu Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly
195 200 205

Gly Gln Asp Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile
210 215 220

Gln Ala Ile Leu Glu
225

<210> 9

<211> 689

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (3)..(482)

<220>

<221> unsure

<222> (105)

<223> any amino acid

<400> 9

ga gaa ata aaa aat gaa ctt ctc caa gca caa aac cct tct gaa ttg 47
 Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu
 i 5 10 15

tac ctg agc tcc aaa aca gaa ttg cag ggt tta att gga cag ttg gat 95
 Tyr Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp
 20 25 30

gag gta agt ntt gaa aaa aac ccc tgc atc cgg gaa gcc agg aga aga 143
 Glu Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg
 35 40 45

gca gtg atc gag gtg caa act ctg atc aca tat att gac ttg aag gag 191
 Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu
 50 55 60

gcc ctt gag aaa aga aag ctg ttt gct tgt gag gag cac cca tcc cat 239
 Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His
 65 70 75

aaa gcc gtc tgg aac gtc ctt gga aac ttg tct gag atc cag gga gaa 287
 Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu
 80 85 90 95

gtt ctt tca ttt gat gga aat cga acc gat aag aac tac atc cgg ctg 335
 Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu
 100 105 110

gaa gag ctg ctc acc aag cag ctg cta gcc ctg gat gct gtt gat ccg 383
 Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro
 115 120 125

cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt 431
 Gln Gly Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu
 130 135 140

gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag 479
 Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu
 145 150 155

tac tgaaatacca gagatctcac ttttgatact gttttgcact tcatatgtgc 532
 Tyr
 160

ttctatgtat agagagctt cagtcattg attatacgt gcatatttca gtcctca 592
 ttatgattga agcaaattct attcagttatc tgctgcttt gatgttgcaa gacaaatatac 652
 attacagcac gttaactttt ccattcggat caaaaaa 689

<210> 10
 <211> 160
 <212> PRT
 <213> Homo sapiens

<400> 10

Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr
 1 5 10 15

Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu
 20 25 30

Val Ser Xaa Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg Ala
 35 40 45

Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala
 50 55 60

Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys
 65 70 75 80

Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val
 85 90 95

Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu
 100 105 110

Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln
 115 120 125

Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala
 130 135 140

Gin Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu Tyr
 145 150 155 160

<210> 11
 <211> 246
 <212> DNA
 <213> *Caenorhabditis elegans*

<400> 11
atgttttcc gcctttcgt tgaaatattt cactttttt tccagtttt tccccatctc 60
gaccgtttt ggttttcga gaaaaccacg ttccaaatca gcgacatctc tcaaattgag 120
atcataggct ttttgaagat tgctcaaatt atgcttctca tattgcatga gcatttgaa 180
gcccggtca tcaaccaaag catttttcc acccatcaca atgattttat cattttttt 240
aaaatt 246

<210> 12
<211> 210
<212> PRT
<213> *Caenorhabditis elegans*

<400> 12
Met Lys Val Asn Val Ser Cys Ser Ser Val Gln Thr Thr Ile Asp Ile
1 5 10 15

Leu Glu Glu Asn Gln Gly Glu Asp Glu Ser Ile Leu Thr Leu Gly Gln
20 25 30

Leu Arg Asp Arg Ile Ala Thr Asp Asn Asp Val Asp Val Glu Thr Met
35 40 45

Lys Leu Leu His Arg Gly Lys Phe Leu Gln Gly Ala Asp Asp Val Ser
50 55 60

Leu Ser Thr Leu Asn Phe Lys Glu Asn Asp Lys Ile Ile Val Met Gly
65 70 75 80

Gly Lys Asn Ala Leu Val Asp Asp Ala Gly Phe Lys Met Leu Met Gln
85 90 95

Tyr Glu Lys His Asn Leu Ser Asn Leu Gln Lys Ala Tyr Asp Leu Asn
100 105 110

Leu Arg Asp Val Ala Asp Leu Glu Arg Gly Phe Leu Glu Lys Pro Lys
115 120 125

Gln Val Glu Met Gly Lys Lys Leu Glu Lys Lys Val Lys Tyr Phe Asn
130 135 140

Glu Glu Ala Glu Arg His Leu Glu Thr Leu Asp Gly Met Asn Ile Ile
145 150 155 160

Thr Glu Thr Thr Pro Glu Asn Gln Ala Lys Arg Asn Arg Glu Lys Arg
 165 170 175

Lys Thr Leu Val Asn Gly Ile Gln Thr Leu Leu Asn Gln Asn Asp Ala
 180 185 190

Leu Leu Arg Arg Leu Gln Glu Tyr Gln Ser Val Leu Asn Gly Asp Ile
 195 200 205

Pro Glu
 210

<210> 13

<211> 1377

<212> DNA

<213> *Caenorhabditis elegans*

<220>

<221> CDS

<222> (1)..(1377)

<400> 13

atg cca gtc gtg aac ata cca atc aaa ata ctt ggt cag aat caa tca 48
 Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser
 1 5 10 15

cat agt cga agt aac tcc tcg tct tct gtt gac aac gat cga aat caa 96
 His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
 20 25 30

cca cca cag cag cca cct caa ccg caa cca caa cag caa tct cag caa 144
 Pro Pro Gln Gln Pro Pro Gln Pro Gln Pro Gln Gln Ser Gln Gln
 35 40 45

caa tac cag cag gct cca aac gtg aat acc aat atg cat cat tcc aac 192
 Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn
 50 55 60

gga ttc tca cct aac ttc cca tct cgt agt cct att ccg gac ttt ccc 240
 Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro
 65 70 75 80

agt ttt tca tct ggg ttc cca aac gat tct gaa tgg tct tcg aat ttc 288
 Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe
 85 90 95

ccg tcg ttt cca aat ttc cca agt gga ttc tca aat gga agt tct aat 336
 Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn
 100 105 110

ttc cct gat ttt cca aga ttc gga aga gat gga gga cta tcg cca aac 384
 Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn
 115 120 125

cca ccg atg caa gga tac agg aga agt cca aca cca aca tca act caa 432
 Pro Pro Met Gln Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gln
 130 135 140

tct cca act tct aca tta aga cgc aac tct cag cag aat caa gct cct 480
 Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gln Ala Pro
 145 150 155 160

cca caa tat tct cag caa caa cca caa caa gct caa caa cgt cag aca 528
 Pro Gln Tyr Ser Gln Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr
 165 170 175

act cct ccg tca aca aaa gct tca tct cga cca cca tct cgt act cgt 576
 Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg
 180 185 190

gaa cca aag gaa cct gag gta ccc gag aga cca gca gtt att cca ttg 624
 Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu
 195 200 205

cca tat gag aag aag gag aaa cca ctg gag aag aaa ggt agt cgt gat 672
 Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp
 210 215 220

tct gga aag ggt gat gag aac ctt gaa gag aac att gcc aag atc acg 720
 Ser Gly Lys Gly Asp Glu Asn Leu Glu Asn Ile Ala Lys Ile Thr
 225 230 235 240

atc gga aag aat aat tgc gag tta tgt ccg gaa caa gaa acg gac ggc 768
 Ile Gly Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly
 245 250 255

gac cca tct cca cta acc tcc cca atc acc gaa gga aag cca aag aga 816
 Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg
 260 265 270

gga aag aaa ctt caa cgt aat caa agt gtt gat ttc aat gcc aag 864
 Gly Lys Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys
 275 280 285

aca att gtt act ttg gat aaa att gaa tta caa gtt gag cag ttg aga 912
 Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg
 290 295 300

aaa aaa gct gct gaa ctc gaa atg gaa aaa gag caa att ctt cgt tct 960
 Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser
 305 310 315 320

cta gga gaa atc agt gtt cat aac tgc atg ttc aaa ctg gaa gaa tgt 1008
 Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys
 325 330 335

gat cgt gaa gag att gaa gca atc act gac cga ttg aca aaa aga aca 1056
 Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr
 340 345 350

aag aca gtt caa gtt gtt gtc gaa act cca cga aat gaa gaa cag aaa 1104
 Lys Thr Val Gln Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys
 355 360 365

aaa gca ctg gaa gat gca act ttg atg atc gat gaa gtc gga gaa atg 1152
 Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met
 370 375 380

atg cat tcg aat att gaa aag gct aag ctg tgc cta caa acc tac atg 1200
 Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met
 385 390 395 400

aac gcc tgt tcg tac gaa gaa act gct gga gcc acc tgc caa aac ttc 1248
 Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe
 405 410 415

ttg aag atc ata att cag tgc gct gct gat gat cag aaa cgc atc aag 1296
 Leu Lys Ile Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
 420 425 430

cgt cgt ctg gaa aat ctg atg tct caa att gag aat gct gag aga acg 1344
 Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr
 435 440 445

aaa gca gat ttg atg gat gat caa agc gaa tag 1377
 Lys Ala Asp Leu Met Asp Asp Gln Ser Glu
 450 455

<210> 14

<211> 458

<212> PRT

<213> *Caenorhabditis elegans*

<400> 14

Met Pro Val Val Asn Ile Pro Ile Lys Ile Leu Gly Gln Asn Gln Ser
1 5 10 15

His Ser Arg Ser Asn Ser Ser Ser Val Asp Asn Asp Arg Asn Gln
20 25 30

Pro Pro Gln Gln Pro Pro Gln Pro Gln Gln Gln Ser Gln Gln
35 40 45

Gln Tyr Gln Gln Ala Pro Asn Val Asn Thr Asn Met His His Ser Asn
50 55 60

Gly Phe Ser Pro Asn Phe Pro Ser Arg Ser Pro Ile Pro Asp Phe Pro
65 70 75 80

Ser Phe Ser Ser Gly Phe Pro Asn Asp Ser Glu Trp Ser Ser Asn Phe
85 90 95

Pro Ser Phe Pro Asn Phe Pro Ser Gly Phe Ser Asn Gly Ser Ser Asn
100 105 110

Phe Pro Asp Phe Pro Arg Phe Gly Arg Asp Gly Gly Leu Ser Pro Asn
115 120 125

Pro Pro Met Gln Gly Tyr Arg Arg Ser Pro Thr Pro Thr Ser Thr Gln
130 135 140

Ser Pro Thr Ser Thr Leu Arg Arg Asn Ser Gln Gln Asn Gln Ala Pro
145 150 155 160

Pro Gln Tyr Ser Gln Gln Pro Gln Gln Ala Gln Gln Arg Gln Thr
165 170 175

Thr Pro Pro Ser Thr Lys Ala Ser Ser Arg Pro Pro Ser Arg Thr Arg
180 185 190

Glu Pro Lys Glu Pro Glu Val Pro Glu Arg Pro Ala Val Ile Pro Leu
195 200 205

Pro Tyr Glu Lys Lys Glu Lys Pro Leu Glu Lys Lys Gly Ser Arg Asp
210 215 220

Ser Gly Lys Gly Asp Glu Asn Leu Glu Glu Asn Ile Ala Lys Ile Thr
225 230 235 240

Ile Gly Lys Asn Asn Cys Glu Leu Cys Pro Glu Gln Glu Thr Asp Gly
245 250 255

Asp Pro Ser Pro Leu Thr Ser Pro Ile Thr Glu Gly Lys Pro Lys Arg
260 265 270

Gly Lys Lys Leu Gln Arg Asn Gln Ser Val Val Asp Phe Asn Ala Lys
275 280 285

Thr Ile Val Thr Leu Asp Lys Ile Glu Leu Gln Val Glu Gln Leu Arg
290 295 300

Lys Lys Ala Ala Glu Leu Glu Met Glu Lys Glu Gln Ile Leu Arg Ser
305 310 315 320

Leu Gly Glu Ile Ser Val His Asn Cys Met Phe Lys Leu Glu Glu Cys
325 330 335

Asp Arg Glu Glu Ile Glu Ala Ile Thr Asp Arg Leu Thr Lys Arg Thr
340 345 350

Lys Thr Val Gln Val Val Glu Thr Pro Arg Asn Glu Glu Gln Lys
355 360 365

Lys Ala Leu Glu Asp Ala Thr Leu Met Ile Asp Glu Val Gly Glu Met
370 375 380

Met His Ser Asn Ile Glu Lys Ala Lys Leu Cys Leu Gln Thr Tyr Met
385 390 395 400

Asn Ala Cys Ser Tyr Glu Glu Thr Ala Gly Ala Thr Cys Gln Asn Phe
405 410 415

Leu Lys Ile Ile Ile Gln Cys Ala Ala Asp Asp Gln Lys Arg Ile Lys
420 425 430

Arg Arg Leu Glu Asn Leu Met Ser Gln Ile Glu Asn Ala Glu Arg Thr
435 440 445

Lys Ala Asp Leu Met Asp Asp Gln Ser Glu
450 455

<210> 15
<211> 588
<212> DNA
<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1)...(588)

<400> 15

atg tca gaa aag act agc aca gtt aca ata cac tat gga aat cag cga 48
 Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Asn Gln Arg
 1 5 10 15

ttt ccg gta gca gtc aat cta aat gag acg tta agt gaa ctg att gat 96
 Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp
 20 25 30

gat tta ctt gaa acg act gag att tct gag aag aaa gtc aag ctt ttt 144
 Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe
 35 40 45

tac gct ggc aag cgt tta aaa gac aaa aaa gcc tcg tta tca aaa ttg 192
 Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu
 50 55 60

ggg tta aaa aat cat agt aaa att cta tgt ata aga cca cat aag caa 240
 Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln
 65 70 75 80

caa cga ggt tcc aag gaa aaa gac acg gtt gag ccc gct ccg aaa gcg 288
 Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala
 85 90 95

gaa gcg gag aat cct gta ttt tcg cgt att tct gga gaa ata aaa gcc 336
 Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala
 100 105 110

atc gat cag tat gtt gac aaa gaa ctt tcc ccc atg tac gac aat tac 384
 Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr
 115 120 125

gta aat aaa ccg tcg aac gat cca aag cag aaa aac aaa cag aaa cta 432
 Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu
 130 135 140

atg ata agt gaa cta ctt tta caa cag ctt tta aaa ttg gat gga gtt 480
 Met Ile Ser Glu Leu Leu Gin Gln Leu Leu Lys Leu Asp Gly Val
 145 150 155 160

gac gta ctg ggc agc gag aaa ttg cgt ttt gaa cgg aag caa ctt gtt 528
 Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val
 165 170 175

tct aag atc caa aaa atg ttg gat cac gtt gac caa aca agc caa gaa 576
Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu
180 185 190

gtg gcc gca tag 588
Val Ala Ala
195

<210> 16
<211> 195
<212> PRT
<213> Schizosaccharomyces pombe

<400> 16
Met Ser Glu Lys Thr Ser Thr Val Thr Ile His Tyr Gly Asn Gln Arg
1 5 10 15

Phe Pro Val Ala Val Asn Leu Asn Glu Thr Leu Ser Glu Leu Ile Asp
20 25 30

Asp Leu Leu Glu Thr Thr Glu Ile Ser Glu Lys Lys Val Lys Leu Phe
35 40 45

Tyr Ala Gly Lys Arg Leu Lys Asp Lys Lys Ala Ser Leu Ser Lys Leu
50 55 60

Gly Leu Lys Asn His Ser Lys Ile Leu Cys Ile Arg Pro His Lys Gln
65 70 75 80

Gln Arg Gly Ser Lys Glu Lys Asp Thr Val Glu Pro Ala Pro Lys Ala
85 90 95

Glu Ala Glu Asn Pro Val Phe Ser Arg Ile Ser Gly Glu Ile Lys Ala
100 105 110

Ile Asp Gln Tyr Val Asp Lys Glu Leu Ser Pro Met Tyr Asp Asn Tyr
115 120 125

Val Asn Lys Pro Ser Asn Asp Pro Lys Gln Lys Asn Lys Gln Lys Leu
130 135 140

Met Ile Ser Glu Leu Leu Gln Gln Leu Leu Lys Leu Asp Gly Val
145 150 155 160

Asp Val Leu Gly Ser Glu Lys Leu Arg Phe Glu Arg Lys Gln Leu Val
165 170 175

Ser Lys Ile Gln Lys Met Leu Asp His Val Asp Gln Thr Ser Gln Glu
 180 185 190

Val Ala Ala
 195

<210> 17

<211> 621

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1)...(621)

<400> 17

atg tct ttt acc cag ttg tgt tct atg gat aaa aaa tat tgg atc 48
 Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile
 1 5 10 15

tct cta gct gta ttg tca gtt act gtt ttg att agc gca tta ttg aaa 96
 Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys
 20 25 30

aag aga gct act gaa acc gaa gat att gtc gtt gtt cat tac gat ggc 144
 Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val Val His Tyr Asp Gly
 35 40 45

gaa aag ttg aat ttt gtg ttg cga caa cca agg ctg aat atg gtt tct 192
 Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser
 50 55 60

tac act agt ttt ctt cgt cgc gtg tgc aac gca ttt tca gta atg ccc 240
 Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro
 65 70 75 80

gac aaa gcg tct ctc aag tta aac ggg gtg acc ctc aag gat ggt tca 288
 Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser
 85 90 95

ctt tcc gac caa aat gtg caa aat gga agt gaa tta gag ctc gaa tta 336
 Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu Glu Leu
 100 105 110

ccc aaa ctg agc ccg gca atg caa caa att gaa gca tat ata gat gag 384
 Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu

115	120	125													
ctt caa cag gat ctc gtc cct aaa att gaa gcc ttc tgc caa tcg tct 432															
Leu	Gln	Gln	Asp	Leu	Val	Pro	Lys	Ile	Glu	Ala	Phe	Cys	Gln	Ser	Ser
130				135					140						
ccc gct tcg gca caa gat gtt caa gat ttg cat aca cgc ctt agt gaa 480															
Pro	Ala	Ser	Ala	Gln	Asp	Val	Gln	Asp	Leu	His	Thr	Arg	Leu	Ser	Glu
145				150					155				160		
aca ttg ttg gct agg atg ata aaa tta gat gct gtt aat gtt gaa gac 528															
Thr	Leu	Leu	Ala	Arg	Met	Ile	Lys	Leu	Asp	Ala	Val	Asn	Val	Glu	Asp
165						170				175					
gac cca gaa gct cgt ctt aaa aga aaa gaa gct att cgt tta tct caa 576															
Asp	Pro	Glu	Ala	Arg	Leu	Lys	Arg	Lys	Glu	Ala	Ile	Arg	Leu	Ser	Gln
180						185				190					
caa tat ttg agt aaa cta gat tcc acc aag aat caa aac aaa tga 621															
Gln	Tyr	Leu	Ser	Lys	Leu	Asp	Ser	Thr	Lys	Asn	Gln	Asn	Lys		
195						200				205					
<210> 18															
<211> 206															
<212> PRT															
<213> Schizosaccharomyces pombe															
<400> 18															
Met Ser Phe Phe Thr Gln Leu Cys Ser Met Asp Lys Lys Tyr Trp Ile															
1	5	10	15												
Ser Leu Ala Val Leu Ser Val Thr Val Leu Ile Ser Ala Leu Leu Lys															
20	25	30													
Lys Arg Ala Thr Glu Thr Glu Asp Ile Val Val Val His Tyr Asp Gly															
35	40	45													
Glu Lys Leu Asn Phe Val Leu Arg Gln Pro Arg Leu Asn Met Val Ser															
50	55	60													
Tyr Thr Ser Phe Leu Arg Arg Val Cys Asn Ala Phe Ser Val Met Pro															
65	70	75	80												
Asp Lys Ala Ser Leu Lys Leu Asn Gly Val Thr Leu Lys Asp Gly Ser															
85	90	95													
Leu Ser Asp Gln Asn Val Gln Asn Gly Ser Glu Leu Glu Leu															

100	105	110
Pro Lys Leu Ser Pro Ala Met Gln Gln Ile Glu Ala Tyr Ile Asp Glu		
115	120	125
Leu Gln Gln Asp Leu Val Pro Lys Ile Glu Ala Phe Cys Gln Ser Ser		
130	135	140
Pro Ala Ser Ala Gln Asp Val Gln Asp Leu His Thr Arg Leu Ser Glu		
145	150	155
160		
Thr Leu Leu Ala Arg Met Ile Lys Leu Asp Ala Val Asn Val Glu Asp		
165	170	175
Asp Pro Glu Ala Arg Leu Lys Arg Lys Glu Ala Ile Arg Leu Ser Gln		
180	185	190
Gln Tyr Leu Ser Lys Leu Asp Ser Thr Lys Asn Gln Asn Lys		
195	200	205

<210> 19
 <211> 2534
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (307)..(2034)

<400> 19
 gcggagctcc gcatccaacc ccggggcccg cgccaaattctt ctggactgga ccagaagttt 60
 ctagccggcc agttgctacc tccctttatc tcctccttcc cctctggcag cgaggaggct 120
 atttccagac acttccaccc ctctctggcc acgtcacccc cgcccttaat tcataaagg 180
 gcccggcgcc ggcttcccg acacgtcggc ggccggagagg ggcccacggc ggcggcccg 240
 ccagagactc ggcgccccga gccagcgccc cgcacccgctt ccccaagcggg cagaccccaa 300
 cccagc atg agc gcc acc cac tcg ccc atg atg cag gtg gcg tcc 348
 Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser
 1 5 10

ggc aac ggt gac cgc gac cct ttg ccc gga tgg gag atc aag atc 396
 Gly Asn Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile
 15 20 25 30

gac ccg cag acc ggc tgg ccc ttc ttc gtg gac cac aac aac agc cgc acc 444
 Asp Pro Gln Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr
 35 40 45

act acg tgg aac gac ccg cgc gtg ccc tct gag ggc ccc aag gag act 492
 Thr Thr Trp Asn Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr
 50 55 60

cca tcc tct gcc aat ggc cct tcc cgg gag ggc tct agg ctg ccg cct 540
 Pro Ser Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro
 65 70 75

gct agg gaa ggc cac cct gtg tac ccc cag ctc cga cca ggc tac att 588
 Ala Arg Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile
 80 85 90

ccc att cct gtg ctc cat gaa ggc gct gag aac cgg cag gtg cac cct 636
 Pro Ile Pro Val Leu His Glu Gly Ala Glu Asn Arg Gln Val His Pro
 95 100 105 110

ttc cat gtc tat ccc cag cct ggg atg cag cga ttc cga act gag gcg 684
 Phe His Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Glu Ala
 115 120 125

gca gca gcg gct cct cag agg tcc cag tca cct ctg cgg ggc atg cca 732
 Ala Ala Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro
 130 135 140

gaa acc act cag cca gat aaa cag tgt gga cag gtc gca gcg gcg gcg 780
 Glu Thr Thr Gln Pro Asp Lys Gln Cys Gly Gln Val Ala Ala Ala Ala
 145 150 155

gca gcc cag ccc cca gcc tcc cac gga cct gag cgg tcc cag tct cca 828
 Ala Ala Gln Pro Pro Ala Ser His Gly Pro Glu Arg Ser Gln Ser Pro
 160 165 170

gct gcc tct gac tca tcc tca tcc tcc tcg gcc agc ctg cct tcc 876
 Ala Ala Ser Asp Cys Ser Ser Ser Ser Ala Ser Leu Pro Ser
 175 180 185 190

tcc ggc agg agc agc ctg ggc agt cac cag ctc ccg cgg ggg tac atc 924
 Ser Gly Arg Ser Ser Leu Gly Ser His Gln Leu Pro Arg Gly Tyr Ile
 195 200 205

tcc att ccg gtg ata cac gag cag aac gtt acc cgg cca gca gcc cag 972
 Ser Ile Pro Val Ile His Glu Gln Asn Val Thr Arg Pro Ala Ala Gln
 210 215 220

ccc tcc ttc cac aaa gcc cag aag acg cac tac cca gcg cag agg ggt 1020
 Pro Ser Phe His Lys Ala Gln Lys Thr His Tyr Pro Ala Gln Arg Gly
 225 230 235

gag tac cag acc cac cag cct gtg tac cac aag atc cag ggg gat gac 1068
 Glu Tyr Gln Thr His Gln Pro Val Tyr His Lys Ile Gln Gly Asp Asp
 240 245 250

tgg gag ccc cgg ccc ctg cgg gcg gca tcc ccg ttc agg tca tct gtc 1116
 Trp Glu Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe Arg Ser Ser Val
 255 260 265 270

cag ggt gca tcg acg cgg gag ggc tca cca gcc agg acg acg cca 1164
 Gln Gly Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg Ser Ser Thr Pro
 275 280 285

ctc cac tcc ccc tcg ccc atc cgt gtg cac acc gtg gtc gac agg cct 1212
 Leu His Ser Pro Ser Pro Ile Arg Val His Thr Val Val Asp Arg Pro
 290 295 300

cag cag ccc atg acc cat cga gaa act gca cct gtt tcc cag cct gaa 1260
 Gln Gln Pro Met Thr His Arg Glu Thr Ala Pro Val Ser Gln Pro Glu
 305 310 315

aac aaa cca gaa agt aag cca ggc cca gtt gga cca gaa ctc cct cct 1308
 Asn Lys Pro Glu Ser Lys Pro Gly Pro Val Gly Pro Glu Leu Pro Pro
 320 325 330

gga cac atc cca att caa gtg atc cgc aaa gag gtg gat tct aaa cct 1356
 Gly His Ile Pro Ile Gln Val Ile Arg Lys Glu Val Asp Ser Lys Pro
 335 340 345 350

gtt tcc cag aag ccc cca cct ccc tct gag aag gta gag gtg aaa gtt 1404
 Val Ser Gln Lys Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val
 355 360 365

ccc cct gct cca gtt cct tgt cct ccc agc cct ggc cct tct gct 1452
 Pro Pro Ala Pro Val Pro Cys Pro Pro Ser Pro Gly Pro Ser Ala
 370 375 380

gtc ccc tct tcc ccc aag agt gtg gct aca gaa gag agg gca gcc ccc 1500
 Val Pro Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro
 385 390 395

agc act gcc cct gca gaa gct aca cct cca aaa cca gga gaa gcc gag 1548
 Ser Thr Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu
 400 405 410

gct ccc cca aaa cat cca gga gtg ctg aaa gtg gaa gcc atc ctg gag 1596
 Ala Pro Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu
 415 420 425 430

aag gtg cag ggg ctg gag cag gct gta gac aac ttt gaa ggc aag aag 1644
 Lys Val Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys
 435 440 445

act gac aaa aag tac ctg atg atc gaa gag tat ttg acc aaa gag ctg 1692
 Thr Asp Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu
 450 455 460

ctg gcc ctg gat tca gtg gac ccc gag gga cga gcc gat gtg cgt cag 1740
 Leu Ala Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln
 465 470 475

gcc agg aga gac ggt gtc agg aag gtt cag acc atc ttg gaa aaa ctt 1785
 Ala Arg Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu
 480 485 490

gaa cag aaa gcc att gat gtc cca ggt caa gtc cag gtc tat gaa ctc 1836
 Glu Gln Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu
 495 500 505 510

cag ccc agc aac ctt gaa gca gat cag cca ctg cag gca atc atg gag 1884
 Gln Pro Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu
 515 520 525

atg ggt gcc gtg gca gca gac aag ggc aag aaa aat gct gga aat gca 1930
 Met Gly Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala
 530 535 540

gaa gat ccc cac aca gaa acc cag cag cca gaa gcc aca gca gca gcg 1980
 Glu Asp Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala
 545 550 555

act tca aac ccc agc agc atg aca gac acc cct ggt aac cca gca gca 2028
 Thr Ser Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala
 560 565 570

ccg tag cctctgcctt gtaaaaatca gactcgaaac cgatgtgtgc ttttagggat 2084
 Pro
 575

tttaagttgc atgcatttca gagactttaa gtcagtttgtt ttttatttgc tgcttggat 2144

gcagtaactt gggtggaggc aaaacactaa taaaagggtt aaaaaggaaa atgatgctt 2204

tcttctataat tcttactctg tacaaataaa gaagttgtt 2264
 ccgttgcgttgc ttctgcagcc ctgtctactt gggcacccccc accacactgtt agctgtggtt 2324
 gtgcactgtc tttttagtct ctggactgga gggtagatg gggagtcaat tacccatcac 2384
 ataaatatga aacatttac agaaatgtt ccattttat gagatgattt tcttcatttc 2444
 ataattaaaa tacctgactt tagagagagt aaaatgtgcc aggagccata ggaatatctg 2504
 tatgttggat gactttaatg ctacatttc 2534

<210> 20
 <211> 575
 <212> PRT
 <213> Homo sapiens

<400> 20
 Met Ser Ala Ala Thr His Ser Pro Met Met Gln Val Ala Ser Gly Asn
 1 5 10 15

Gly Asp Arg Asp Pro Leu Pro Pro Gly Trp Glu Ile Lys Ile Asp Pro
 20 25 30

Gln Thr Gly Trp Pro Phe Phe Val Asp His Asn Ser Arg Thr Thr Thr
 35 40 45

Trp Asn Asp Pro Arg Val Pro Ser Glu Gly Pro Lys Glu Thr Pro Ser
 50 55 60

Ser Ala Asn Gly Pro Ser Arg Glu Gly Ser Arg Leu Pro Pro Ala Arg
 65 70 75 80

Glu Gly His Pro Val Tyr Pro Gln Leu Arg Pro Gly Tyr Ile Pro Ile
 85 90 95

Pro Val Leu His Glu Gly Ala Glu Asn Arg Gln Val His Pro Phe His
 100 105 110

Val Tyr Pro Gln Pro Gly Met Gln Arg Phe Arg Thr Gln Ala Ala Ala
 115 120 125

Ala Ala Pro Gln Arg Ser Gln Ser Pro Leu Arg Gly Met Pro Glu Thr
 130 135 140

Thr Gln Pro Asp Lys Gln Cys Gly Gln Val Ala Ala Ala Ala Ala

145 150 155 160

Gln Pro Pro Ala Ser His Gly Pro Glu Arg Ser Gln Ser Pro Ala Ala
165 170 175

Ser Asp Cys Ser Ser Ser Ser Ala Ser Leu Pro Ser Ser Gly
180 185 190

Arg Ser Ser Leu Gly Ser His Gln Leu Pro Arg Gly Tyr Ile Ser Ile
195 200 205

Pro Val Ile His Glu Gln Asn Val Thr Arg Pro Ala Ala Gln Pro Ser
210 215 220

Phe His Lys Ala Gln Lys Thr His Tyr Pro Ala Gln Arg Gly Glu Tyr
225 230 235 240

Gln Thr His Gln Pro Val Tyr His Lys Ile Gln Gly Asp Asp Trp Glu
245 250 255

Pro Arg Pro Leu Arg Ala Ala Ser Pro Phe Arg Ser Ser Val Gln Gly
260 265 270

Ala Ser Ser Arg Glu Gly Ser Pro Ala Arg Ser Ser Thr Pro Leu His
275 280 285

Ser Pro Ser Pro Ile Arg Val His Thr Val Val Asp Arg Pro Gln Gln
290 295 300

Pro Met Thr His Arg Glu Thr Ala Pro Val Ser Gln Pro Glu Asn Lys
305 310 315 320

Pro Glu Ser Lys Pro Gly Pro Val Gly Pro Glu Leu Pro Pro Gly His
325 330 335

Ile Pro Ile Gln Val Ile Arg Lys Glu Val Asp Ser Lys Pro Val Ser
340 345 350

Gln Lys Pro Pro Pro Ser Glu Lys Val Glu Val Lys Val Pro Pro
355 360 365

Ala Pro Val Pro Cys Pro Pro Ser Pro Gly Pro Ser Ala Val Pro
370 375 380

Ser Ser Pro Lys Ser Val Ala Thr Glu Glu Arg Ala Ala Pro Ser Thr
385 390 395 400

Ala Pro Ala Glu Ala Thr Pro Pro Lys Pro Gly Glu Ala Glu Ala Pro

405

410

415

Pro Lys His Pro Gly Val Leu Lys Val Glu Ala Ile Leu Glu Lys Val
420 425 430

Gln Gly Leu Glu Gln Ala Val Asp Asn Phe Glu Gly Lys Lys Thr Asp
435 440 445

Lys Lys Tyr Leu Met Ile Glu Glu Tyr Leu Thr Lys Glu Leu Leu Ala
450 455 460

Leu Asp Ser Val Asp Pro Glu Gly Arg Ala Asp Val Arg Gln Ala Arg
465 470 475 480

Arg Asp Gly Val Arg Lys Val Gln Thr Ile Leu Glu Lys Leu Glu Gln
485 490 495

Lys Ala Ile Asp Val Pro Gly Gln Val Gln Val Tyr Glu Leu Gln Pro
500 505 510

Ser Asn Leu Glu Ala Asp Gln Pro Leu Gln Ala Ile Met Glu Met Gly
515 520 525

Ala Val Ala Ala Asp Lys Gly Lys Lys Asn Ala Gly Asn Ala Glu Asp
530 535 540

Pro His Thr Glu Thr Gln Gln Pro Glu Ala Thr Ala Ala Ala Thr Ser
545 550 555 560

Asn Pro Ser Ser Met Thr Asp Thr Pro Gly Asn Pro Ala Ala Pro
565 570 575

<210> 21

<211> 1966

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (43)..(1416)

<400> 21

cggtggagc ggggcgggaa gcgcgttcagg gcagcggatc cc atg tcg gcc ctg 54
Met Ser Ala Leu
1

agg cgc tcg ggc tac ggc ccc agt gac ggt ccg tcc tac ggc cgc tac 102

Arg	Arg	Ser	Gly	Tyr	Gly	Pro	Ser	Asp	Gly	Pro	Ser	Tyr	Gly	Arg	Tyr	
5						10				15				20		
tac ggg cct ggg ggt gga gat gtg ccg gta cac cca cct cca ccc tta															150	
Tyr	Gly	Pro	Gly	Gly	Asp	Val	Pro	Val	His	Pro	Pro	Pro	Leu			
						25				30				35		
tat cct ctt cgc cct gaa cct ccc cag cct ccc att tcc tgg cgg gtg															198	
Tyr	Pro	Leu	Arg	Pro	Glu	Pro	Pro	Gln	Pro	Pro	Ile	Ser	Trp	Arg	Val	
						40				45				50		
cgc ggg ggc ggc ccg gcg gag acc acc tgg ctg gga gaa ggc gga gga															246	
Arg	Gly	Gly	Gly	Pro	Ala	Glu	Thr	Thr	Trp	Leu	Gly	Glu	Gly	Gly		
						55				60				65		
ggc gat ggc tac tat ccc tcg gga ggc gcc tgg cca gag cct ggt cga															294	
Gly	Asp	Gly	Tyr	Tyr	Pro	Ser	Gly	Gly	Ala	Trp	Pro	Glu	Pro	Gly	Arg	
						70				75				80		
gcc gga gga agc cac cag gag cag cca cca tat cct agc tac aat tct															342	
Ala	Gly	Gly	Ser	His	Gln	Glu	Gln	Pro	Pro	Tyr	Pro	Ser	Tyr	Asn	Ser	
						85				90				100		
aac tat tgg aat tct act gcg aga tct agg gct cct tac cca agt aca															390	
Asn	Tyr	Trp	Asn	Ser	Thr	Ala	Arg	Ser	Arg	Ala	Pro	Tyr	Pro	Ser	Thr	
						105				110				115		
tat cct gta aga cca gaa ttg caa ggc cag agt ttg aat tct tat aca															438	
Tyr	Pro	Val	Arg	Pro	Glu	Leu	Gln	Gly	Gln	Ser	Leu	Asn	Ser	Tyr	Thr	
						120				125				130		
aat gga gcg tat ggt cca aca tac ccc cca ggc cct ggg gca aat act															486	
Asn	Gly	Ala	Tyr	Gly	Pro	Thr	Tyr	Pro	Pro	Gly	Pro	Gly	Ala	Asn	Thr	
						135				140				145		
gcc tca tac tca ggg gct tat tat gca cct ggt tat act cag acc agt															534	
Ala	Ser	Tyr	Ser	Gly	Ala	Tyr	Tyr	Ala	Pro	Gly	Tyr	Thr	Gln	Thr	Ser	
						150				155				160		
tac tcc aca gaa gtt cca agt act tac cgt tca tct ggc aac agc cca															582	
Tyr	Ser	Thr	Glu	Val	Pro	Ser	Thr	Tyr	Arg	Ser	Ser	Gly	Asn	Ser	Pro	
						165				170				175		
act cca gtc tct cgt tgg atc tat ccc cag cag gac tgt cag act gaa															630	
Thr	Pro	Val	Ser	Arg	Trp	Ile	Tyr	Pro	Gln	Gln	Asp	Cys	Gln	Thr	Glu	
						185				190				195		
gca ccc cct ctt agg ggg cag gtt cca gga tat ccg cct tca cag aac															678	

Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro Pro Ser Gln Asn
 200 205 210

cct gga atg acc ctg ccc cat tat cct tat gga gat ggt aat cgt agt 726
 Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp Gly Asn Arg Ser
 215 220 225

gtt cca caa tca gga ccg act gta cga cca caa gaa gat gcg tgg gct 774
 Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu Asp Ala Trp Ala
 230 235 240

tct cct ggt gct tat gga atg ggt ggc cgt tat ccc tgg cct tca tca 822
 Ser Pro Gly Ala Tyr Gly Met Gly Arg Tyr Pro Trp Pro Ser Ser
 245 250 255 260

gcg ccc tca gca cca ccc ggc aat ctc tac atg act gaa agt act tca 870
 Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr Glu Ser Thr Ser
 265 270 275

cca tgg cct agc agt ggc tct ccc cag tca ccc cct tca ccc cca gtc 918
 Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro Ser Pro Pro Val
 280 285 290

cag cag ccc aag gat tct tca tac ccc tat agc caa tca gat caa agc 966
 Gln Gln Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln Ser Asp Gln Ser
 295 300 305

atg aac cgg cac aac ttt cct tgc agt gtc cat cag tac gaa tcc tcg 1014
 Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln Tyr Glu Ser Ser
 310 315 320

ggg aca gtg atc aat gaa gat tca gat ctt ttg gat tcc caa gtc cag 1062
 Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp Ser Gln Val Gln
 325 330 335 340

tat agt gct gag cct cag ctg tat ggt aat gcc acc agt gac cat ccc 1110
 Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr Ser Asp His Pro
 345 350 355

aac aat caa gat caa agt agc agt ctt cct gaa gaa tgt gta cct tca 1158
 Asn Asn Gln Asp Gln Ser Ser Leu Pro Glu Glu Cys Val Pro Ser
 360 365 370

gat gaa agt act cct ccg agt att aaa aaa atc ata cat gtg ctg gag 1206
 Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile His Val Leu Glu
 375 380 385

aag gtc cag tat ctt gaa caa gaa gta gaa gaa ttt gta gga aaa aag 1254

Lys Val Gin Tyr Leu Glu Gln Glu Val Glu Glu Phe Val Gly Lys Lys
 390 395 400

aca gac aaa gca tac tgg ctt ctg gaa gaa aig cta acc aag gaa ctt 1302
 Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu Thr Lys Glu Leu
 405 410 415 420

ttg gaa ctg gat tca gtt gaa act ggg ggc cag gac tct gta cgg cag 1350
 Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp Ser Val Arg Gln
 425 430 435

gcc aga aaa gag gct gtt tgt aag att cag gcc ata ctg gaa aaa tta 1398
 Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile Leu Glu Lys Leu
 440 445 450

gaa aaa aaa gga tta tga aaggat tag aacaaagtgg aagcctgtta 1446
 Glu Lys Lys Gly Leu
 455

ctaacttgac caaagaacac ttgatttagt taattaccct cttttgaaa tgcctgtga 1506

tgacaagaag caatacattc cagctttcc tttgattta tacttgaaaaa actggcaaag 1566

gaatggaaga atattttagt catgaagttg tttcagttt tcagacgaat gaatgtaata 1626

ggaaactatg gagttaccaa tattgccaag tagactcaact ccttaaaaaa tttatggata 1686

tctacaagct gcttattacc agcaggaggg aaacacactt cacacaacag gcttatcaga 1746

aacctaccag atgaaactgg atataattt agacaaacag gatgtgtttt tttaaacatc 1806

tggatatctt gtcacatttt tgtacattgt gactgcttcc aacatataact tcatgtgtaa 1866

ttatagctta gacttagcc ttcttgact tctgtttgt ttgttattt gcagttaca 1926

aatatagtat tattctctaa aaaaaaaaaa aaaaaaaaaa 1966

<210> 22
 <211> 457
 <212> PRT
 <213> Homo sapiens

<400> 22
 Met Ser Ala Leu Arg Arg Ser Gly Tyr Gly Pro Ser Asp Gly Pro Ser
 1 5 10 15

Tyr Gly Arg Tyr Tyr Gly Pro Gly Gly Asp Val Pro Val His Pro

20	25	30
Pro Pro Pro Leu Tyr Pro Leu Arg Pro Glu Pro Pro Gln Pro Pro Ile		
35	40	45
Ser Trp Arg Val Arg Gly Gly Pro Ala Glu Thr Thr Trp Leu Gly		
50	55	60
Glu Gly Gly Gly Asp Gly Tyr Tyr Pro Ser Gly Gly Ala Trp Pro		
65	70	75
80		
Glu Pro Gly Arg Ala Gly Gly Ser His Gln Glu Gln Pro Pro Tyr Pro		
85	90	95
Ser Tyr Asn Ser Asn Tyr Trp Asn Ser Thr Ala Arg Ser Arg Ala Pro		
100	105	110
Tyr Pro Ser Thr Tyr Pro Val Arg Pro Glu Leu Gln Gly Gln Ser Leu		
115	120	125
Asn Ser Tyr Thr Asn Gly Ala Tyr Gly Pro Thr Tyr Pro Pro Gly Pro		
130	135	140
Gly Ala Asn Thr Ala Ser Tyr Ser Gly Ala Tyr Tyr Ala Pro Gly Tyr		
145	150	155
160		
Thr Gln Thr Ser Tyr Ser Thr Glu Val Pro Ser Thr Tyr Arg Ser Ser		
165	170	175
Gly Asn Ser Pro Thr Pro Val Ser Arg Trp Ile Tyr Pro Gln Gln Asp		
180	185	190
Cys Gln Thr Glu Ala Pro Pro Leu Arg Gly Gln Val Pro Gly Tyr Pro		
195	200	205
Pro Ser Gln Asn Pro Gly Met Thr Leu Pro His Tyr Pro Tyr Gly Asp		
210	215	220
Gly Asn Arg Ser Val Pro Gln Ser Gly Pro Thr Val Arg Pro Gln Glu		
225	230	235
240		
Asp Ala Trp Ala Ser Pro Gly Ala Tyr Gly Met Gly Gly Arg Tyr Pro		
245	250	255
Trp Pro Ser Ser Ala Pro Ser Ala Pro Pro Gly Asn Leu Tyr Met Thr		
260	265	270
Glu Ser Thr Ser Pro Trp Pro Ser Ser Gly Ser Pro Gln Ser Pro Pro		

275	280	285
Ser Pro Pro Val Gln Gln Pro Lys Asp Ser Ser Tyr Pro Tyr Ser Gln		
290	295	300
Ser Asp Gln Ser Met Asn Arg His Asn Phe Pro Cys Ser Val His Gln		
305	310	315
Tyr Glu Ser Ser Gly Thr Val Ile Asn Glu Asp Ser Asp Leu Leu Asp		
325	330	335
Ser Gln Val Gln Tyr Ser Ala Glu Pro Gln Leu Tyr Gly Asn Ala Thr		
340	345	350
Ser Asp His Pro Asn Asn Gln Asp Gln Ser Ser Ser Leu Pro Glu Glu		
355	360	365
Cys Val Pro Ser Asp Glu Ser Thr Pro Pro Ser Ile Lys Lys Ile Ile		
370	375	380
His Val Leu Glu Lys Val Gln Tyr Leu Glu Gln Glu Val Glu Glu Phe		
385	390	395
Val Gly Lys Lys Thr Asp Lys Ala Tyr Trp Leu Leu Glu Glu Met Leu		
405	410	415
Thr Lys Glu Leu Leu Glu Leu Asp Ser Val Glu Thr Gly Gly Gln Asp		
420	425	430
Ser Val Arg Gln Ala Arg Lys Glu Ala Val Cys Lys Ile Gln Ala Ile		
435	440	445
Leu Glu Lys Leu Glu Lys Lys Gly Leu		
450	455	
<210> 23		
<211> 4308		
<212> DNA		
<213> Homo sapiens		
<220>		
<221> CDS		
<222> (247)...(1590)		
<400> 23		
cccccccccc cccccccccc ccngaagacg cccggagcgg ctgctgcagc cagtagcggc 60		

cccttcacccg gctgcccccgc tcagacctag tcggggagggg tgcgaggcat gcagctgggg 120
 gcccagctcc ggtgccgcac cccgtaaaagg gctgatctc cacctcgcca cctcagccac 180
 gggacgccaa gaccgcattcc aattcagact tctttggtg cttgtgaaac tgaacacaac 240
 aaaagt atg gat atg gga aac caa cat cct tct att agt agg ctt cag 288
 Met Asp Met Gly Asn Gln His Pro Ser Ile Ser Arg Leu Gln
 1 5 10
 gaa atc caa aag gaa gta aaa agt gta gaa cag caa gtt atc ggc ttc 336
 Glu Ile Gln Lys Glu Val Lys Ser Val Glu Gln Gln Val Ile Gly Phe
 15 20 25 30
 agt ggt ctg tca gat gac aag aat tac aag aaa ctg gag agg att cta 384
 Ser Gly Leu Ser Asp Asp Lys Asn Tyr Lys Lys Leu Glu Arg Ile Leu
 35 40 45
 aca aaa cag ctt ttt gaa ata gac tct gta gat act gaa gga aaa gga 432
 Thr Lys Gln Leu Phe Glu Ile Asp Ser Val Asp Thr Glu Gly Lys Gly
 50 55 60
 gat att cag caa gct agg aag cgg gca gca cag gag aca gaa cgt ctt 480
 Asp Ile Gln Gln Ala Arg Lys Arg Ala Ala Gln Glu Thr Glu Arg Leu
 65 70 75
 ctc aaa gag ttg gag cag aat gca aac cac cca cac cgg att gaa ata 528
 Leu Lys Glu Leu Glu Gln Asn Ala Asn His Pro His Arg Ile Glu Ile
 80 85 90
 cag aac att ttt gag gaa gcc cag tcc ctc gtg aga gag aaa att gtg 576
 Gln Asn Ile Phe Glu Glu Ala Gln Ser Leu Val Arg Glu Lys Ile Val
 95 100 105 110
 cca ttt tat aat gga ggc aac tgc gta act gat gag ttt gaa gaa ggc 624
 Pro Phe Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly
 115 120 125
 atc caa gat atc att ctg agg ctg aca cat gtt aaa act gga gga aaa 672
 Ile Gln Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys
 130 135 140
 atc tcc ttg cgg aaa gca agg tat cac act tta acc aaa atc tgt gcg 720
 Ile Ser Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala
 145 150 155
 gtg caa gag ata atc gaa gac tgc atg aaa aag cag cct tcc ctg ccg 768
 Val Gln Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro

160	165	170	
ctt tcc gag gat gca cat cct tcc gtt gcc aaa atc aac ttc gtg atg 816			
Leu Ser Glu Asp Ala His Pro Ser Val Ala Lys Ile Asn Phe Val Met			
175	180	185	190
tgc gag gtg aac aag gcc cga ggg gtc ctg att gca ctt ctg atg ggt 864			
Cys Glu Val Asn Lys Ala Arg Gly Val Leu Ile Ala Leu Leu Met Gly			
195	200	205	
gtg aac aac aat gag acc tgc agg cac tta tcc tgt gtg ctc tcg ggg 912			
Val Asn Asn Asn Glu Thr Cys Arg His Leu Ser Cys Val Leu Ser Gly			
210	215	220	
ctg atc gct gac ctg gat gct cta gat gtg tgc ggc cgg aca gaa atc 960			
Leu Ile Ala Asp Leu Asp Ala Leu Asp Val Cys Gly Arg Thr Glu Ile			
225	230	235	
aga aat tat cgg agg gag gta gta gaa gat atc aac aaa tta ttg aaa 1008			
Arg Asn Tyr Arg Arg Glu Val Val Glu Asp Ile Asn Lys Leu Leu Lys			
240	245	250	
tat ctg gat ttg gaa gag gaa gca gac aca act aaa gca ttt gac ctg 1056			
Tyr Leu Asp Leu Glu Glu Ala Asp Thr Thr Lys Ala Phe Asp Leu			
255	260	265	270
aga cag aat cat tcc att tta aaa ata gaa aag gtc ctc aag aga atg 1104			
Arg Gln Asn His Ser Ile Leu Lys Ile Glu Lys Val Leu Lys Arg Met			
275	280	285	
aga gaa ata aaa aat gaa ctt ctc caa gca caa aac cct tct gaa ttg 1152			
Arg Glu Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu			
290	295	300	
tac ctg agc tcc aaa aca gaa ttg cag ggt tta att gga cag ttg gat 1200			
Tyr Leu Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp			
305	310	315	
gag gta agt ctt gaa aaa aac ccc tgc atc cgg gaa gcc agg aga aga 1248			
Glu Val Ser Leu Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg			
320	325	330	
gca gtg atc gag gtg caa act ctg atc aca tat att gac ttg aag gag 1296			
Ala Val Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu			
335	340	345	350
gcc ctt gag aaa aga aag ctg ttt gct tgt gag gag cac cca tcc cat 1344			
Ala Leu Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His			

355	360	365	
aaa gcc gtc tgg aac gtc ctt gga aac ttg tct gag atc cag gga gaa			1392
Lys Ala Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gin Gly Glu			
370	375	380	
gtt ctt tca ttt gat gga aat cga acc gat aag aac tac atc cgg ctg			1440
Val Leu Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu			
385	390	395	
gaa gag ctg ctc acc aag cag ctg cta gcc ctg gat gct gat gat ccg			1488
Glu Glu Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro			
400	405	410	
cag gga gaa gag aag tgt aag gct gcc agg aaa caa gct gtg agg ctt			1536
Gln Gly Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu			
415	420	425	430
gcg cag aat att ctc agc tat ctc gac ctg aaa tct gat gaa tgg gag			1584
Ala Gln Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu			
435	440	445	
tac tga aataccagag atctcacttt tgatactgtt ttgeacttca tatgtgcttc			1640
Tyr			
tatgtataga gagcttcag ttcatggatt tatacgtgca tatttcagtc tcagtttta			1700
tgattgaagc aaattctatt cagtatctgc tgctttgtat gttgcaagac aaatatcatt			1760
acagcacgtt aactttcca ttccggatcat tatctgtatg atgtgggtgt gtttgggtgg			1820
tttgccttt ttttgcgtt ttaatcaga aaacaaaata gaggcagctt ttgttagattt			1880
taaatgggtt gtgcaagcat taaaatgcag gtctttcaga atctagaact aggcataacc			1940
ttacataata ctagaaaaat tatgagaaag gggaaatttt tggtaaata agagtaaggt			2000
tcaaacacaa gcagtacatg ttctgtttca ttatgctgca tagaaggctt tttttcact			2060
tataaggcct gattggcctt acccagctt acggggtggg gttttttgt ttgttcagac			2120
agtctgtct ttgtaaaca ttttagttt gaaaaacagc atctgcattt tccccatcct			2180
ctacgtttta gagaggaatc ttgtttttgt gtgcaacata agaaaattat gaaaactaat			2240
agccaaaaaa cctttgagat tgcattaaag agaaggata aaggaccagc aataatacct			2300
tgttaagttgc ttttgtttgt aaaatctgag cttatagtt tccttagtga gtaaattcat			2360

aaggatggga acatttaat taagttaatg ggcccttaaa aaaaaaaaaag gaaacactca 2420
tacctgttgt gggaggatga atactggaga cgggttacca atgtcaggtt atactaaaac 2480
taaatcagaa agtctgaatg tagcacataa tggttctctt ctgttgtcca aggctgtaaa 2540
atggacagcc ttgtcacacc tccccggtgc tggtttacaa cgtgagggtt gacgctgtca 2600
gtaacccaga gggaccaggc ctcccttaggt tttctaggca gtcagctgtt aaccactcac 2660
ttagtaaatg tcataactac acctgctcca ggaccaatca gtgaaacctg ctcggaatta 2720
aaggcttcct ctgggtgcct gctgaacaac tgagctcatg tcatggcat gtgggtttt 2780
ctctgttgcc tgaaagagcc attaaagtca gtcgtgcgtg aagcatctct cttctaaagg 2840
atgtgtatcc ccataaatgc ttctgagga tccggtacaa aatgatttcc caaagttctg 2900
aagtgccttg agaacatgtg ggtcccgagtg ttataacaga ctcctccccc gggtcacctt 2960
ttgcctggtc atcctgttag agtacatctt tgaaatcca gggtaatatt ctcttcaga 3020
gatgctcatt gtgttaactct gtgttagggag atagtcactt taaacagctc aaagtagcta 3080
gctaaaggag tagccttaaa tacctaaaag atgacagaag catagccctt aacaaatctt 3140
cagcttgtct ctcagtatcc cccaatcatg aaaatccctt gctatgtctt tcctactaga 3200
aatgttctag aatcgctgga cggtgggtc agagggcagt cggtatccat tagcgtgagct 3260
tcccatacta ctgcagggtcc aactcctggc aaccgcgggc tcaaggcagg tcattggat 3320
ccacgtttg gccacagtag ttgttaggatt gctttctgt atcataatccat tagaatgctc 3380
ttaaaaatctt gaggaagagt ttttattttt tatttatttt tgagatggag tctctgttg 3440
ccaggctgca gtgcagtggt gccatctcag ctcactgcaa cctccaccc tcaggttcaa 3500
gcgattctcc tgcctcagcc acctgagtag ctgggagtagt acggcatgtgg caccatgcct 3560
ggctaatttt tgcattttt atagagttga gatttcacca tgatggcag gctggctcg 3620
aactcctgac ctcgtgatcc gccccccctcg gccccccaaa gtgctggat taacgggtgt 3680
gagccacggc gcccagccca ggaagagttt ttaaatttgc gctctgttta attataaccac 3740
tggaaaatca tggttacgct tcaggcataat tcttccccag agtactactt acatttaaa 3800

tttcattttg taaaagttaaa tgcgtcaggcatt ccctttaaaa gtgtccattg ttctttgaaa 3860
 gtagacgtt cagtcattct tttcaaaacaa gtgtttgtgt acctttgcc aagctgtggg 3920
 catcggtgtgt gagtacaggg tgctcagctc ttccaccgtc attttgaatt gttcacatgg 3980
 gtaattggtc atggaaatga tcagattgac ctgttgcac tgtcaggcat ggctttgttt 4040
 ctatgttcaa tctgttctcg ttccttgcac cggattattc tactcctgca atgaaccctg 4100
 ttgacaccgg attttagctct tgtcggccctt cgtggggagc tgtttgcgtt aatatgagct 4160
 actgcgtatgtt attcttaaac tgggcttgac acattgtatt gtattttgcgtt gatctgtat 4220
 gaaaagaatc tgtactgcaa gtaaaaccta ctccccaaaa atgtgtggct ttgggtctgc 4280
 attaaacgct gtagtccatg ttcatgcc 4308

<210> 24

<211> 447

<212> PRT

<213> Homo sapiens

<400> 24

Met	Asp	Met	Gly	Asn	Gln	His	Pro	Ser	Ile	Ser	Arg	Leu	Gln	Glu	Ile
1									10					15	

Gln	Lys	Glu	Val	Lys	Ser	Val	Glu	Gln	Gln	Val	Ile	Gly	Phe	Ser	Gly
									25				30		

Leu	Ser	Asp	Asp	Lys	Asn	Tyr	Lys	Lys	Leu	Glu	Arg	Ile	Leu	Thr	Lys
								35				40		45	

Gln	Leu	Phe	Glu	Ile	Asp	Ser	Val	Asp	Thr	Glu	Gly	Lys	Gly	Asp	Ile
								50				55		60	

Gln	Gln	Ala	Arg	Lys	Arg	Ala	Ala	Gln	Glu	Thr	Glu	Arg	Leu	Leu	Lys
								65				70		75	

Glu	Leu	Glu	Gln	Asn	Ala	Asn	His	Pro	His	Arg	Ile	Glu	Ile	Gln	Asn
								85			90			95	

Ile	Phe	Glu	Glu	Ala	Gln	Ser	Leu	Val	Arg	Glu	Lys	Ile	Val	Pro	Phe
								100				105		110	

Tyr Asn Gly Gly Asn Cys Val Thr Asp Glu Phe Glu Glu Gly Ile Gln

115	120	125
Asp Ile Ile Leu Arg Leu Thr His Val Lys Thr Gly Gly Lys Ile Ser		
130	135	140
Leu Arg Lys Ala Arg Tyr His Thr Leu Thr Lys Ile Cys Ala Val Gln		
145	150	155
160		
Glu Ile Ile Glu Asp Cys Met Lys Lys Gln Pro Ser Leu Pro Leu Ser		
165	170	175
180		
Glu Asp Ala His Pro Ser Val Ala Lys Ile Asn Phe Val Met Cys Glu		
185	190	
195		
Val Asn Lys Ala Arg Gly Val Leu Ile Ala Leu Leu Met Gly Val Asn		
200	205	
210		
Asn Asn Glu Thr Cys Arg His Leu Ser Cys Val Leu Ser Gly Leu Ile		
215	220	
225		
Ala Asp Leu Asp Ala Leu Asp Val Cys Gly Arg Thr Glu Ile Arg Asn		
230	235	240
245		
Tyr Arg Arg Glu Val Val Glu Asp Ile Asn Lys Leu Leu Lys Tyr Leu		
250	255	
260		
Asp Leu Glu Glu Ala Asp Thr Thr Lys Ala Phe Asp Leu Arg Gln		
265	270	
275		
Asn His Ser Ile Leu Lys Ile Glu Lys Val Leu Lys Arg Met Arg Glu		
280	285	
290		
Ile Lys Asn Glu Leu Leu Gln Ala Gln Asn Pro Ser Glu Leu Tyr Leu		
295	300	
305		
Ser Ser Lys Thr Glu Leu Gln Gly Leu Ile Gly Gln Leu Asp Glu Val		
310	315	320
325		
Ser Leu Glu Lys Asn Pro Cys Ile Arg Glu Ala Arg Arg Arg Ala Val		
330	335	
340		
Ile Glu Val Gln Thr Leu Ile Thr Tyr Ile Asp Leu Lys Glu Ala Leu		
345	350	
355		
Glu Lys Arg Lys Leu Phe Ala Cys Glu Glu His Pro Ser His Lys Ala		
360	365	
Val Trp Asn Val Leu Gly Asn Leu Ser Glu Ile Gln Gly Glu Val Leu		

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/21053

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.. 1, 13, 24, 25
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

No meaningful search could be carried out because no limitations could be placed on the sequence.

3. Claims Nos.
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/21053

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database Genbank, National Center for Biotech. Info., Accession No. G29287, MYERS, R.M., 04 October 1996, see entire reference.	2,4
X	Database Genbank, National Center for Biotech. Info., Accession No. G06974, HUDSON, T., "Whitehead Institute/MIT Center for Genome Research,'19 October 1995, see entire reference.	2,4
X	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. V81267, OTSUKA PHARM CO LTD, 'New Bcl-2 interaction protein gene (Bis)- useful for elucidation of the molecular mechanism of apoptosis, and in diagnosis, prevention and treatment of diseases,' 15 December 1998 see entire reference.	2-5
X	Database, Geneseq, Derwent, Alexandria, Virginia, Accession No. T19051, MATSUBARA ET AL., "Identifying gene signatures in 3'-directed human cDNA library,' 01 June 1995, see entire reference.	2,4
X	Database Geneseq, Derwent, Alexandria, Virginia, Accession No. Q90296, LA JOLLA CANCER RES FOUN. 'Human Bcl-2-associated protein BAG-1 cDNA,'18 May 1995 see entire reference.	2-5,14

370	375	380
Ser Phe Asp Gly Asn Arg Thr Asp Lys Asn Tyr Ile Arg Leu Glu Glu		
385	390	395
Leu Leu Thr Lys Gln Leu Leu Ala Leu Asp Ala Val Asp Pro Gln Gly		
405	410	415
Glu Glu Lys Cys Lys Ala Ala Arg Lys Gln Ala Val Arg Leu Ala Gln		
420	425	430
Asn Ile Leu Ser Tyr Leu Asp Leu Lys Ser Asp Glu Trp Glu Tyr		
435	440	445

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/21053

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :07N 21/02; C07K 1/00

US CL :530/387.1, 350; 435/6, 7/1; 536/23.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/387.1, 350; 435/6, 7/1; 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,652,223 A (KOHN ET AL) 29 July 1997(29/7/97) see entire document.	2-5, 14, 32-34
X	Database Genbank-EST, National Center for Biotech. Info., Accession No. AA693697, HILLIER, L. ET AL. 'WashU-NCI human EST Project,' 16 December 1997, see entire reference.	2
X	Database Genbank-EST, National Center for Biotech. Info., Accession No. AA456862, NCI CGAP, 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index,' 15 August 1997, see entire reference.	2,4

 Further documents are listed in the continuation of Box C.

See patent family annex.

• Special categories of cited documents:	*T*	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance		
B earlier document published on or after the international filing date	*X*	document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
O document referring to an oral disclosure, use, exhibition or other means	*&*	document member of the same patent family
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

24 NOVEMBER 1999

Date of mailing of the international search report

19 JAN 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

SHEELA J. HUFF

Facsimile No. (703) 305-3230

Telephone No. (703) 308-0196